Analysis of Error Introduced During End-User Post-Processing of Airborne Laser Data (LiDAR)

Sarah Louise Smith

A thesis submitted for the degree of Doctor of Philosophy

Department of Geography

University College London

ProQuest LLC
789 East Eisenhower Parkway
P.O. Box 1346
Ann Arbor, MI 48106-1346

Abstract

The primary aims and objectives of this thesis are to identify the sources and operation of the errors which are introduced during end-user post-processing of airborne laser scanning data. Previous research has concentrated on the errors incorporated during data capture and preliminary supplier processing. The errors which are introduced by the end-users have been largely neglected. As a result, data users cannot currently estimate the errors within, and therefore the quality of, the models they produce.

Laser scanning is a remote sensing technique for the capture of height data of the surface of the Earth. It offers competitive capture costs, high accuracy, and is particularly suited to capturing information in complex urban areas. As a result the commercial value of laser scanning data is high. However, in order to realise the potential of this technique, the quality of the datasets derived from the data must be assessed and the errors introduced during modelling understood. For users to make informed decisions regarding the design of their post-processing workflow it is fundamental that they know how and where errors may be introduced. The characteristics of these errors are investigated in this thesis using a range of approaches.

End-user post-processing is divided into three techniques in the thesis: data structuring, filtering and segmentation. Each process is investigated in terms of accuracy and sensitivity, through the comparison of several methods with reference models. New algorithms for filtering and segmenting laser data are presented. The errors created by each process are identified and analysed. The location of errors across the elevation surface are also investigated. It is shown how this information could be used to aid end-users design their post-processing methodology. The methdology for analysing the errors is presented as a framework which could be used as a standard for ALS models.

This thesis shows that the choice of post-processing methodology can significantly alter both the magnitude and spatial pattern of errors with a model derived from airborne laser scanning data. The differences between modelling strategies, and the importance of these differences, is shown with reference to a flood modelling application. Finally, strategies for minimising error for post-processing are proposed.

Declaration

I declare that the work within this thesis is my own, unless otherwise stated. It has not been submitted for any other degree.

Contents

Table of contents 2

List of figures 15

List of tables 34

Acknowledgements 38

1 Introduction **40**

1.1 Background and the Choice of Airborne Laser Scanning 42

1.2 Post-Processing of Airborne Laser Scanning Data 45

1.2.1 Data Structuring 45

1.2.2 Data Filtering 48

1.2.3 Segmentation 48

1.3 Modelling the Errors in Urban ALS Models 50

1.3.1 Definition of Terms 50

1.3.2 Methods for Modelling Error 52

1.4 Motivation For This Research 53

1.4.1 Problem 53

1.4.2 Drivers 54

1.5 Aim of the Thesis 55

1.6 Thesis Outline 57

1.7 Chapter Summary 59

2 The Theory of Airborne Laser Scanning **60**

2.1 Chapter Introduction 60

2.2 Laser Theory 60

2.3 Laser Scanning for Remote Sensing of the Earth 62

2.4 Topographic Airborne Laser Scanning 65

2.4.1 System Components 65

2.5 Characteristics of the Raw Data 68

2.5.1 Intensity of Recorded Pulse 68

2.5.2 Return Attributes 71

2.5.3 Sources of Error in Data Capture 73

2.6 Supplier Processing 75

2.7 Chapter Summary 83

3 Research Context **84**

3.1 Chapter Introduction 84

3.2 Representations of Surfaces 85

3.2.1 Elevation Modelling 88

3.3 Urban Remote Sensing 91

3.4 Previous Modelling Techniques Using Airborne Laser Scanning Data 96

3.4.1 Filtering Strategies 96

3.4.2 Segmentation Techniques 101

3.5 Measuring Error in Elevation Modelling 102

3.5.1 Error in Discrete Models 104

3.5.2 Error in Continuous Models 111

3.5.3 Significance of Errors in Elevation Models 112

3.6 Chapter Summary 114

4 Data and Methodology **115**

4.1 Chapter Overview 115

4.2 Aims of the methodology 115

4.3 Data and Study Sites 117

4.3.1 Airborne Laser Datasets 117

4.4 Reference data 122

4.4.1 Terrestrial Laser Data 124

4.4.2 Photogrammetric DTM 128

4.4.3 RTK GPS Readings 129

4.5 Methodology 130

4.5.1 Error in Data Capture and Preliminary Post Processing 130

4.5.2 Data Restructuring Investigation 131

4.5.2.1 Nearest neighbour interpolation 136

4.5.2.2 Bilinear interpolation 136

4.5.2.3 Bicubic interpolation 138

4.5.2.4 Biharmonic Splining 139

4.5.2.5 Geostatistical Interpolation: Kriging 141

4.5.3 Quantifying and Analysing the Error in Re-structuring 145

4.5.3.1 Split-sample Testing 146

4.5.3.2 Cross validation 148

4.5.3.3 Jackknifing 149

4.5.3.4 Validation of the Kriged Surfaces 150

4.5.4 Data Filtering Investigation 152

4.5.5 The Effect of the Segmentation Process 157

4.6 Assessing the Adequacy of the Methodology 161

4.7 Chapter Summary 163

5 Quantifying the Data Capture Errors **165**

5.1 Chapter Overview and Aims 165

5.2 Quantifying the Error in Data Capture 165

5.3 Chapter Summary 170

6 Pilot Study **171**

6.1 Chapter Introduction 171

6.2 Background to Investigation 172

6.3 Pilot Study Methodology 173

6.3.1 Results from Pilot Study Split Sample Test 181

6.3.2 Results from Cross Validation 189

6.3.3 Results from Jackknifing 190

6.3.4 Results from Kriging Variance 192

6.3.5 Summary of Pilot Study Findings 193

6.4 Chapter Summary 195

7 Analysing the Errors Introduced during Data Restructuring **196**

7.1 Chapter Introduction 196

7.2 Main Investigation 196

7.2.1 Study Site 196

7.2.2 Analysing the Height Errors 212

7.2.3 Summary of Findings from Main Investigation 213

7.3 Significance of the Findings for Urban Flood Modelling and Other Real World Applications 213

7.4 Conclusions 220

7.5 Chapter Summary 221

8 Analysis of the Error Arising Through Filtering 223

8.1 Chapter Introduction 223

8.2 Chapter Aims 224

8.3 The Filtering Algorithms 225

8.3.1 The Grid Based Geometric Classification Filtering Technique 226

8.3.2 Iterative Robust Interpolation (IRI) Filtering 239

8.3.3 Adaptive TIN Models (ATM) 244

8.4 Quantitative Comparison of the Filtered Datasets: Classification Accuracy 252

8.4.1 Results and Discussion: Overall classification accuracy 259

8.4.1.1 The IRI Method 263

8.4.1.2 The ATM Method 264

8.4.1.3 The Geometric Method 264

8.4.2 Results and Discussion: Producer's Accuracy 265

8.4.3 Type I vs Type II Classification Errors 267

8.5 Quantitative Comparison of Filtered Datasets: Accuracy of the Height Modelling . 275

8.5.1 Results and Discussion: Image Differencing 276

8.5.1.1 Statistical Analysis 280

8.5.2 Results and Discussion: Comparison with GCPs 284

8.6 Answering the Research Questions 288

8.6.1 Q1 Are there significant differences between the classification accuracies for different filtering methods? 288

8.6.2 Q2 Is there a significant difference in accuracy caused by filter method parameter changes? 290

8.6.3 Q3 Is there a significant difference in the classification accuracy between the gridded datasets and the raw data? 292

8.6.4 Q4 Are there significant differences between the classification accuracies of the different gridded datasets? 292

8.6.5 Q5 Does the spatial resolution of the data affect the characteristics of the error in the filtered datasets? 296

8.7 Conclusions and the Significance of the Findings 298

8.8 Chapter Summary 299

9 Analysis of the Error Arising Through Segmentation 300

9.1 Chapter Introduction 300

9.2 Chapter Aims 301

9.3 Segmentation methods 301

9.3.1 Geometric Segmentation 302

9.3.2 ATM Segmentation 303

9.4 Assessing the Success of the Segmentation 306

9.4.1 Classification Accuracy 306

9.4.2 Types of Error 312

9.4.3 Accuracy of Height Modelling 318

9.5 Answering the Research Questions 322

9.5.1 Q1 Does using different segmentation algorithms cause a significant difference in accuracy? 322

9.5.2 Q2 Does changing the parameter settings for the segmentation algorithms cause a significant difference in accuracy? 323

9.5.3 Q3 Does the use of raw or gridded data cause a significant difference in segmentation accuracy? 323

9.5.4 Q4 Does the use of different data restructuring methods cause a significant difference in segmentation accuracy? 324

9.5.5 Q5 Does decreasing the spatial resolution of the dataset cause significant differences in segmentation accuracy? 325

9.5.6 Q6 Are there significant differences between the classification accuracies for each of the classes within the datasets? 326

9.6 Summary of Findings and Discussion of Wider Relevance 327

9.7 Chapter Summary 327

10 Discussion and Visualisation of Results **329**

10.1 Chapter Overview and Aims 329

10.2 Summary of Principal Research Findings 329

10.2.1 Restructuring errors 330

10.2.2 Filtering Errors 331

10.2.3 Segmentation Errors 332

10.3 How could this information be used? 332

10.3.1 Example 1 335

10.3.2 Example 2 335

10.3.3 Example 3 337

10.3.4 Discussion 338

10.4 How can this error information be communicated? 339

10.5 Chapter Summary 345

11 Conclusions **346**

11.1 Contribution to the Research Literature 346

11.2 Applicability of the Results 347

11.3 Recommendations 348

11.4 Further Research 348

11.4.1 Error propagation 348

11.4.2 Visualising and understanding the spatial patterns of errors 349

11.4.3 Repeat testing 350

11.4.4 Additional extensions 350

11.5 Final Word 351

Appendices **351**

A List of Publications From Thesis **352**

A.1 Published Articles 352

A.2 Recently Submitted Articles 353

B Code and Functions **354**

B.1 Code for interpolation of laser data onto regular grids at different resolutions . . . 354

B.2 RMSE function 355

B.3 Subset funtion 355

B.4 Cross validation routine 356

B.5 Function used for randomly spliting dataset 358

B.6 Exclude funtion used for cross validation routine 359

C Workflow for Matching Individual Scans **360**

C.1 Characteristics of the Raw Data 360

C.2 Workflow in Polyworks IMAlign 360

D Digital Surface Models **362**

D.1 Digital Surfaces Models Created Using Different Restructuring Techniques 362

E Scripts for Geometric Method **365**

E.1 Binary Classification 365

E.2 Example script for cookie-cutting the image 366

E.3 Script for removing the object halos 368

E.4 Script for interpolating across the gaps to create the continuous surface 369

F Example Parameter File for IRI Method **375**

G The resultant DTMs from filtering using IRI algorithm **386**

G.1 IRI Default method 386

G.2 IRI Adapted method 392

H Resultant DTMs and Associated Settings from ATM pilot study **398**

H.1 Settings 1 - Terrain angle set at 5 degrees 398

H.1.1 Run 1 398

H.1.2 Run 2 398

H.1.3 Run 3 399

H.1.4 Run 4 399

H.2 Settings 2 - Terrain angle set at 10 degrees 403

H.2.1 Run 1 403

H.2.2 Run 2 403

H.2.3 Run 3 403

H.2.4 Run 4 403

H.3 Settings 3 - Terrain angle set at 15 degrees 407

H.3.1 Run 1 407

H.3.2 Run 2 407

H.3.3 Run 3 407

H.3.4 Run 4 407

H.4 Settings 4 - Terrain angle set at 20 degrees 411

H.4.1 Run 1 411

H.4.2 Run 2 411

H.4.3 Run 3 411

H.4.4 Run 4 411

I Results from Repeat Experiments Using Four Research Assistants 415

I.1 Comparison with Research Assistant 1 416

I.2 Comparison with Research Assistant 2 416

I.3 Comparison with Research Assistant 3 417

I.4 Comparison with Research Assistant 4 417

J Producer's Accuracy Results 418

K t-Test Results 421

L Workings for Research Questions 1-4: Chapter 8 426

L.1 Q1 Are there significant differences between the classification accuracies for different filtering methods? 426

L.2 Q2 Is there a significant difference in accuracy caused by filter method parameter changes? 431

L.3 Q3 Is there a significant difference in the classification accuracy between the gridded datasets and the raw data? 432

L.4 Q4 Are there significant differences in in the classification accuracies of the different gridded datasets? 433

M Workings for Research Question 5: Chapter 8 **435**

N Results and Workings for Research Question 2: Chapter 9 **438**

N.1 Comparison of default and adapted settings for Geometric algorithm 438

N.2 Comparison of default and adapted settings for Geometric algorithm 440

O Results and Workings for Research Question 4: Chapter 9 **441**

P Results and Workings for Research Question 6: Chapter 9 **446**

List of Figures

1.1 Workflow diagram showing the processing of laser data from capture to user application. Red circle indicates principal focus for thesis research 41

1.2 Flow diagram showing stages in the post-processing of Airborne Laser Scanner data 46

2.1 Elements of laser sensor (after Measures 1992) 63

2.2 Airborne system functional block diagram 65

2.3 Airborne laser scanning 68

2.4 Laser data coloured by height (m), here displayed as an image 69

2.5 Laser data of subset of study area. Data coloured by intensity 70

2.6 Typical Range Measurement Scenario (after Lindenberger et al 2003) 72

2.7 Airborne Laser Scanning Reference Systems (after Lindenberger et al 2003) 78

2.8 Modelling from Laser Data 82

3.1 Diagram depicting research context for thesis, and outline of issues discussed in this chapter 85

4.1 Orthorectified aerial photograph of the pilot study region showing area 80m by 50m. The church can be seen on the centre left of the image, distinct clumps of mature vegetation are clearly seen in the central region. On the right-hand side there are some industrial flat roofed buildings 119

4.2 Map showing location of main study area in local context. Study area shown by circle. Scale of mapping 1:1 000 000. Reproduced from Ordnance Survey map data by permission of Ordnance Survey, © Crown copyright 120

4.3 Map showing zoomed in location of the main study area, the boundary of which is shown by the solid red line. Scale of mapping 1: 25 000. Reproduced from Ordnance Survey map data by permission of Ordnance Survey, © Crown copyright 121

4.4 Orthorectified aerial photograph of study site ($1km^2$) 122

4.5 Contours in the study region derived from a linear surface model (1m resolution) . 122

4.6 Shaded surface for 1m linear model derived from laser data 123

4.7 Dimensional drawing of the LMS-Z420i terrestrial laser scanning instrument. Diagram reproduced with kind permission of Dr Graham Hunter of Riegl 125

4.8 Components of the equipment: at the base is the Riegl LMS-Z420i scanner, mounted on top is the Nikon camera, with the GPS antenna above 126

4.9 View of the terrestrial lidar dataset. Terrestrial laser points are coloured by RGB values attributed from the digital photograph captured at the same time as laser scanning 127

4.10 View of DTM created using photogrammetric techniques. Note the remaining footprints of many of the buildings, and the obvious smoothing of some areas 129

4.11 Example of bilinear interpolation 138

4.12 Locations of ground control points at which RTK GPS readings were captured. Locations are marked by crosshairs 157

4.13 Example normalised Digital Surface Model 159

4.14 Showing the measurement of object heights within the reference dataset: example 1. Note the arrow shows the measurement from the base of the building (ground height) to the edge of the building roof 161

4.15 Showing the measurement of object heights within the reference dataset: example 2 161

5.1 3D Point Cloud for the Scan with no Colour Attributed 166

5.2 3D Point Cloud for the Scan. Colour for the points attributed from associated imagery 167

5.3 Terrestrial and airborne laser points. Terrestrial points shown in brown shades, airborne points shown in green. Note the close general correspondence between the two datasets 167

6.1 The original laser points for the pilot study displayed as a point cloud 174

6.2 Surface interpolated using bilinear method onto a 1m grid 175

6.3 Surface interpolated using bilinear method onto a 1m grid with original laser points overlain 175

6.4 Surface interpolated using bicubic method onto a 1m grid 176

6.5 Surface interpolated using bicubic method onto a 1m grid with original laser points overlain 176

6.6 Surface interpolated using nearest neighbour method onto a 1m grid 177

6.7 Surface interpolated using nearest neighbour method onto a 1m grid with original laser points overlain 177

6.8 Surface interpolated using biharmonic spline method onto a 1m grid 178

6.9 Surface interpolated using biharmonic spline method onto a 1m grid with original laser points overlain 178

6.10 Semivarigram computed using all of the raw data point 180

6.11 Oblique view of surface interpolated using ordinary block kriging method onto a 1m grid 180

6.12 Oblique view of surface interpolated using ordinary block kriging method onto a 1m grid 181

6.13 Oblique view of surface interpolated using bilinear method on a 1m grid 182

6.14 Oblique view of surface interpolated using bilinear method on a 2m grid 182

6.15 Oblique view of surface interpolated using bilinear method on a 4m grid 183

6.16 Spatial Pattern of Errors on the Bicubic 1m Surface 187

6.17 Spatial Pattern of Errors on the Biharmonic Spline 1m Surface 187

6.18 Observed vs. predicted height values using the bilinear interpolation method at 1m grid resolution for the 'leave one out' cross validation method 191

6.19 Observed vs. predicted height values using the nearest neighbour interpolation method at 1m grid resolution for the 'leave one out' cross validation method 191

6.20 Map of Kriging Standard Error for 1m Interpolation Grid. 193

7.1 Orthorectified aerial photograph of the 1km^2 Bristol study site 197

7.2 View of the hill shaded surface created using the bilinear algorithm at 1m resolution 199

7.3 View of the hill shaded surface created using the bicubic algorithm at 1m resolution 199

7.4 View of the hill shaded surface created using the nearest neighbour algorithm at 1m resolution 200

7.5 Surface profile of the nearest neighbour 1m surface 200

7.6 Example of area dominated by spikes in cubic surface. The locations of spikes can be seen by the 'freckling' of the surface. Darker pixels indicate negative spikes, whilst the lighter 'freckles' indicate positive elevation spikes. Extent of area shown is approximately 225m by 235m 203

7.7 Example of area dominated by spikes in cubic surface. Extent of area shown is approximately 225m by 235m 204

7.8 View of the surface created using the bicubic algorithm at 1m resolution. Note the spikes across the surface. 204

7.9 Histogram showing the distribution of residual values for the bilinear 1m surface. Histogram shows all values, with no outliers removed 206

7.10 Histogram showing the distribution of residual values for the bicubic 1m surface. The histogram shows the distrubution of residuals between -25 and +25m, this excludes the outliers. The full range of residuals for the bicubic surface was between -431.02m and 931.72m, with 165 outliers below the -25m threshold and some 128 outliers above +25m 207

7.11 Histogram showing the distribution of residual values for the nearest neighbour 1m surface. Histogram shows the distribution of residuals between -25m and +25m, excluding two outliers, one which lay at -25.42m and one which lay at 25.68m . . . 208

7.12 Scatter plot showing location and magnitude of errors in bilinear 0.5m surface for the whole Bristol 1km^2 study area 209

7.13 Scatter plot showing location and magnitude of errors over 2m in bilinear 0.5m surface209

7.14 Scatter plot showing location and magnitude of errors over 4m in bilinear 0.5m surface210

7.15 Scatter plot showing location and magnitude of errors over 6m in bilinear 0.5m surface210

7.16 Showing the intensity of the green band, displayed in greyscale 211

7.17 Showing the intensity of green band of original image. Scale bar shows green intensity211

7.18 Resultant classification of aerial photograph 212

7.19 Flooding of cubic 0.5m surface with water level of 12m 215

7.20 Flooding of cubic 0.5m surface with water level of 16m 215

7.21 Flooding of cubic 0.5m surface with water level of 20m 216

7.22 Flooding of cubic 0.5m surface with water level of 24m 216

7.23 Flooding of nearest neighbour 0.5m surface with water level of 12m 216

7.24 Flooding of nearest neighbour 0.5m surface with water level of 16m 216

7.25 Flooding of nearest neighbour 0.5m surface with water level of 20m 218

7.26 Flooding of nearest neighbour 0.5m surface with water level of 24m 218

7.27 4m Flood Event Across Bilinear 1m Surface 218

7.28 5m Flood Event Across Bilinear 1m Surface 218

7.29 4m Flood Event Across Bicubic 1m Surface 219

7.30 5m Flood Event Across Bicubic 1m Surface 219

7.31 4m Flood Event Across Nearest Neighbour 1m Surface 219

7.32 5m Flood Event Across Nearest Neighbour 1m Surface 219

7.33 4m Flood Event Across Biharmonic Spline 1m Surface 219

7.34 5m Flood Event Across Biharmonic Spline 1m Surface 219

7.35 4m Flood Event Across Ordinary Kriging Spline 1m Surface 219

7.36 5m Flood Event Across Biharmonic Spline 1m Surface 219

8.1 Greyscale height image of the bilinear 2m surface before image division for the whole Bristol study area 227

8.2 Showing the divisions of the image created from default parameter settings for the whole Bristol study area 227

8.3 Showing the outline of one of the identified objects for the whole Bristol study area 227

8.4 Showing all of the image divisions in the whole of the Bristol study region 227

8.5 Showing the results of the calculation of object compactness for each of the image divisions 230

8.6 Showing the results of the calculation of the length to width ratio for each image division 230

8.7 Results of calculation showing object width values 230

8.8 Results of calculation of mean object heights 230

8.9 Results of calculation showing mean difference between objects 232

8.10 Resultant binary classification, with above surface objects shown in blue 232

8.11 Gaps produced by the removal of above ground objects in the cookie-cutting stage. Note the 'halo' of remaining pixels around gaps 232

8.12 Following dilation of the above ground areas the halos have been removed before interpolation over the gaps is performed 232

8.13 The continuous DTM created using the bilinear 2m data 233

8.14 Histogram equalisation stretch of DTM in Figure 8.13 233

8.15 DTM created using default method for bilinear 1m data 233

8.16 DTM created using default method for bilinear 0.5m data 233

8.17 DTM created using default method for the nearest neighbour (NN) 2m data for the Bristol study area ($1km^2$) 234

8.18 DTM created using default method for the NN 1m data for the Bristol study area. Lighter pixels represent higher elevation 234

8.19 DTM created using default method for NN 0.5m data 235

8.20 DTM created using default method for bicubic 2m data 235

8.21 DTM created using default method for bicubic 1m data 235

8.22 DTM created using default method for bicubic 0.5m data 235

8.23 DTM created using adapted method for bilinear 2m data 236

8.24 DTM created using adapted method for bilinear 1m data 236

8.25 DTM created using adapted method for bilinear 0.5m data 236

8.26 DTM created using adapted method for NN 2m data 236

8.27 DTM created using adapted method for NN 1m data 237

8.28 DTM created using adapted method for NN 0.5m data 237

8.29 DTM created using adapted method for bicubic 2m data 237

8.30 DTM created using adapted method for bicubic 1m data 237

8.31 DTM created using adapted method for bicubic 0.5m data 238

8.32 Hillshaded image of subset of DTM created using original method 242

8.33 Hillshaded subset of DTM created using adapted method 242

8.34 DTM created using raw input data (with original point spacing) with the default IRI method 243

8.35 DTM created using bilinearly gridded input data at 0.5m resolution with the default IRI method 243

8.36 DTM created using bicubic gridded input data at 0.5m resolution with the default IRI method 244

8.37 DTM created using nearest neighbour gridded input data at 0.5m resolution with the default IRI method 244

8.38 The role of the iteration angle and distance in the ATM algorithm 246

8.39 Terrain/ground points only. Coloured by elevation (m) 248

8.40 Resultant DTM created from ground points. Key in (m) 248

8.41 DTM created from filtering of raw data using default ATM settings 250

8.42 DTM created from Bilinear data at 1m using default ATM method 250

8.43 DTM created from Bicubic interpolated data at 1m using default ATM method . . 250

8.44 DTM created from Nearest Neighbour data at 1m using default ATM method . . . 250

8.45 DTM created from Bilinear data at 0.5m using default ATM method 251

8.46 DTM created from Bilinear data at 2m using default ATM method 251

8.47 An orthorectified aerial photograph of the study area, with one of the filtered point datasets overlain for purposes of comparison of classification. The point dataset shown (pink crosshairs) is that created from the SCOP DTM method for the bicubic 1m data 253

8.48 Original surface (DSM) before filtering of above ground points 257

8.49 DTM created from filtered bicubic 1m surface 257

8.50 DTM created from filtered nearest neighbour 1m surface 257

8.51 Bar chart showing the calculated producer's acuracies for all three filtering methods 266

8.52 Aerial photograph of study area overlain with classification errors. Red crosshairs (Type II) are errors in ground class, blue crosshairs (Type I) show errors in above ground classification for the ATM algorithm. Note the predominance of Type I errors268

8.53 Aerial photograph of study area overlain with classification errors. Red crosshairs (Type II) are errors in ground class, blue crosshairs (Type I) show errors in above ground classification for the geometric algorithm 269

8.54 Aerial photograph of study area overlain with classification errors. Red crosshairs (Type II) are errors in ground class, blue crosshairs (Type I) show errors in above ground classification for the IRI algorithm 270

8.55 Showing the error in the bare earth classification on the multi-roofed building. Note the mis-classified point (red crosshair) in the centre-left of the image. This point is clearly positioned over of part of a building but has been erroneously classified as bare earth by the algorithm 271

8.56 Percentages of total error 274

8.57 Percentages of Type I error 274

8.58 Percentages of Type II error 274

8.59 The filtered geometric linear 2m surface 277

8.60 The photogrammetrically derived reference model 277

8.61 The difference image for the geometric linear 2m dataset. Created by subtracting the Geometric filtered bilinear (2m) dataset from the photogrammetric reference DTM. The height differences in the image range from -14.87 to + 86.65 (m)(Figure 8.60) 277

8.62 The difference image for the ATM linear 2m dataset 278

8.63 The difference image for the IRI cubic 2m dataset 278

8.64 Profile of the filtered cubic 0.5m dataset using the IRI algorithm. Note the presence of some above ground objects remaining within the surface 280

8.65 Profile of the reference DTM. Note the much smoother surface than that produced by the IRI algorithm, with almost no above ground features remaining 281

8.66 Zoomed in section of profile of the filtered cubic 0.5m dataset. Note the irregular surface form and the scouring in built regions 281

8.67 Zoomed in section of profile of the reference DTM. Note the much smoother surface in which no scouring is evident 281

9.1 Showing the relative values of the image objects (for the whole Bristol 1 km^2 study area) for the mean difference to neighbours calculation. Values range between -7 and 16.36m: the darker the pixel the lower the mean difference to neighbours value 303

9.2 Showing the membership function for the bare earth classification. Values defining the threshold ranged between 2 and -10. The left-hand axis indicated the level of fuzziness. For example, the part of the range which is fuzzy lies approximately between -10 and -5. 304

9.3 Showing the membership function for the building classification. The form of the slope indicates that lower values (between 2 and 5) are more fuzzy and are less likely to be classified definitely as building. 304

9.4 Showing the relative values of the image objects for the elliptical fit calculation. The calculated values range between 0.09 and 0.9, the darker the pixel the lower the elliptical fit value 305

9.5 Showing the membership function for the vegetation classification. The form of the slope indicates that values of approximately 1 should be classified as vegetation, but those values greater or less than 1 are much less likely. The limits of the range of acceptable vegetation values are 0.5-1.5, beyond this range image objects will not be considered to be vegetation. 305

9.6 Showing producer's accuracy results for the segmentation of the linear dataset with the adapted geometric algorithm 308

9.7 Showing producer's accuracy results for the segmentation of the linear dataset with the adapted ATM algorithm 308

9.8 Aerial photograph of study area overlain with classification errors for the ATM cubic 1m dataset. Red crosshairs are errors in ground class, blue crosshairs show errors in building class, green crosshairs show errors in vegetation class. 313

9.9 Aerial photograph of study area overlain with classification errors for the geometric linear 2m dataset. Red crosshairs are errors in ground class, blue crosshairs show errors in building class, green crosshairs show errors in vegetation class. 314

9.10 Zoomed in image of incorrectly classified vegetation points (green crosshairs) for the adapted geometric algorithm (linear 2m dataset) 315

9.11 Zoomed in image of incorrectly classified ground points (red crosshairs) for the adapted geometric algorithm (linear 2m dataset) 315

9.12 Zoomed in image of incorrectly classified building points (blue crosshairs) for the adapted geometric algorithm (linear 2m dataset) 316

9.13 Example normalised Digital Surface Model for the nearest neighbour 1m dataset created using the default IRI method. The image shows the relative changes in heights above ground surface: the lighter the pixel the higher it is above ground. . 319

9.14 Showing the measurement of object heights within the reference dataset: example 1 321

10.1 Showing the range of choices the user must make during post-processing of ALS data. For each of the boxes resulting from the restructuring stage, there are the full range of options indicated for filtering. These filtering choices are shown only once (for simplicity) for one of the boxes. The same is true for the segmentation choices shown, with each relating the results from each of the filtering options. 334

10.2 Showing recommended path for user wishing to create a DSM incorporating the least overall error 336

10.3 Showing recommended path for user wishing to create a DTM incorporating the least overall error 336

10.4 Showing recommended path for user wishing to create a segmented dataset which incorporates the least overall error 337

10.5 Plan view of surface created using linear interpolation. X and Y coordinates in metres340

10.6 Plan view of 'uncertainty' surface for the DSM created using bilinear interpolation (1m dataset). For legend see figure 10.7 341

10.7 Plan view of 'uncertainty' surface for the DSM created using bicubic interpolation (1m dataset). Note that legend denotes magnitudes of errors (m) multiplied by 10 for display 342

10.8 Plan view of uncertainty surface for the DSM created using bilinear interpolation (1m dataset) for the main study area. Amounts of probable error displayed are in metres 343

10.9 Plan view of the NW quadrant of the uncertainty surface for the DSM created using bilinear interpolation (1m dataset) for the main study area. Amounts of probable error displayed are in metres 344

D.1 Bilinear 0.5m DSM 362

D.2 Bilinear 1m DSM 362

D.3 Bilinear 2m DSM 363

D.4 Bicubic 0.5m DSM 363

D.5 Bicubic 1m DSM 363

D.6 Bicubic 2m DSM 363

D.7 Nearest Neighbour 0.5m DSM 364

D.8 Nearest Neighbour 1m DSM 364

D.9 Nearest Neighbour 2m DSM 364

G.1 Hill shaded DTM created using raw input data at original point spacing resolution 386

G.2 Height coloured DTM created using raw input data at original point spacing . . . 386

G.3 Contour map of DTM created using raw input data at original point spacing resolution387

G.4 Hillshaded DTM created using input data at bilinear gridded data at 0.5m resolution387

G.5 Height coloured DTM created using input data at bilinear gridded data at 0.5m resolution 387

G.6 Contour map of DTM created using input data at bilinear gridded data at 0.5m resolution 387

G.7 Hillshaded DTM created using bilinear input gridded data at 1m resolution 387

G.8 Height coloured DTM created using input data at bilinear gridded data at 1m resolution 387

G.9 Contour map of DTM created using bilinear input gridded data at 1m resolution . 388

G.10 Hillshaded DTM created using input data at bilinear gridded data at 2m resolution 388

G.11 Height coloured DTM created using input data at bilinear gridded data at 2m resolution 388

G.12 Contour map of DTM created using input data at bilinear gridded data at 2m resolution 388

G.13 Hillshaded DTM created using bicubic input gridded data at 0.5m resolution . . . 388

G.14 Height coloured DTM created using bicubic input gridded data at 0.5m resolution 388

G.15 Contour map of DTM created using bicubic input gridded data at 0.5m resolution 389

G.16 Hillshaded DTM created using bicubic input gridded data at 1m resolution 389

G.17 Height coloured DTM created using bicubic input gridded data at 1m resolution . 389

G.18 Contour map of DTM created using bicubic input gridded data at 1m resolution . 389

G.19 Hillshaded DTM created using bicubic input gridded data at 2m resolution 389

G.20 Height coloured DTM created using bicubic input gridded data at 2m resolution . 389

G.21 Contour map of DTM created using bicubic input gridded data at 2m resolution . 390

G.22 Hillshaded DTM created using nearest neighbour input gridded data at 0.5m resolution390

G.23 Height coloured DTM created using nearest neighbour input gridded data at 0.5m resolution 390

G.24 Contour map of DTM created using nearest neighbour input gridded data at 0.5m resolution 390

G.25 Hillshaded DTM created using nearest neighbour input gridded data at 1m resolution390

G.26 Height coloured DTM created using nearest neighbour input gridded data at 1m resolution 391

G.27 Contour map of DTM created using nearest neighbour input gridded data at 1m resolution 391

G.28 Hillshaded DTM created using nearest neighbour input gridded data at 2m resolution391

G.29 Height coloured DTM created using nearest neighbour input gridded data at 2m resolution 391

G.30 Contour map of DTM created using nearest neighbour input gridded data at 2m resolution 391

G.31 Hill shaded DTM created using raw input data at original point spacing resolution 392

G.32 Height coloured DTM created using raw input data at original point spacing resolution392

G.33 Contour map of DTM created using raw input data at original point spacing resolution392

G.34 Hillshaded DTM created using input data at bilinear gridded data at 0.5m resolution392

G.35 Height coloured DTM created using input data at bilinear gridded data at 0.5m resolution 393

G.36 Contour map of DTM created using input data at bilinear gridded data at 0.5m resolution 393

G.37 Hillshaded DTM created using bilinear input gridded data at 1m resolution 393

G.38 Height coloured DTM created using input data at bilinear gridded data at 1m resolution 393

G.39 Contour map of DTM created using bilinear input gridded data at 1m resolution . 393

G.40 Hillshaded DTM created using input data at bilinear gridded data at 2m resolution 393

G.41 Height coloured DTM created using input data at bilinear gridded data at 2m resolution 394

G.42 Contour map of DTM created using input data at bilinear gridded data at 2m resolution 394

G.43 Hillshaded DTM created using bicubic input gridded data at 0.5m resolution . . . 394

G.44 Height coloured DTM created using bicubic input gridded data at 0.5m resolution 394

G.45 Contour map of DTM created using bicubic input gridded data at 0.5m resolution 394

G.46 Hillshaded DTM created using bicubic input gridded data at 1m resolution 394

G.47 Height coloured DTM created using bicubic input gridded data at 1m resolution . 395

G.48 Contour map of DTM created using bicubic input gridded data at 1m resolution . 395

G.49 Hillshaded DTM created using bicubic input gridded data at 2m resolution 395

G.50 Contour map of DTM created using bicubic input gridded data at 2m resolution . 395

G.51 Height coloured DTM created using bicubic input gridded data at 2m resolution . 395

G.52 Hillshaded DTM created using nearest neighbour input gridded data at 0.5m resolution396

G.53 Height coloured DTM created using nearest neighbour input gridded data at 0.5m resolution 396

G.54 Contour map of DTM created using nearest neighbour input gridded data at 0.5m resolution 396

G.55 Hillshaded DTM created using nearest neighbour input gridded data at 1m resolution396

G.56 Height coloured DTM created using nearest neighbour input gridded data at 1m resolution 396

G.57 Contour map of DTM created using nearest neighbour input gridded data at 1m resolution 396

G.58 Hillshaded DTM created using nearest neighbour input gridded data at 2m resolution397

G.59 Height coloured DTM created using nearest neighbour input gridded data at 2m resolution 397

G.60 Contour map of DTM created using nearest neighbour input gridded data at 2m resolution 397

H.1 Classification created using Settings 1 Run 1 399

H.2 Terrain points created using Settings 1 Run 1 399

H.3 DTM created using Settings 1 Run 2 400

H.4 Classification created using Settings 1 Run 2 400

H.5 Terrain points created using Settings 1 Run 2 400

H.6 DTM created using Settings 1 Run 2 400

H.7 Classification created using Settings 1 Run 3 401

H.8 Terrain points created using Settings 1 Run 3 401

H.9 DTM created using Settings 1 Run 3 401

H.10 Classification created using Settings 1 Run 3 401

H.11 Terrain points created using Settings 1 Run 3 401

H.12 DTM created using Settings 1 Run 3 402

H.13 Classification created using Settings 2 Run 1 404

H.14 Terrain points created using Settings 2 Run 1 404

H.15 DTM created using Settings 2 Run 2 404

H.16 Classification created using Settings 2 Run 2 404

H.17 Terrain points created using Settings 2 Run 2 404

H.18 DTM created using Settings 2 Run 2 405

H.19 Classification created using Settings 2 Run 3 405

H.20 Terrain points created using Settings 2 Run 3 405

H.21 DTM created using Settings 2 Run 3 405

H.22 Classification created using Settings 2 Run 3 406

H.23 Terrain points created using Settings 2 Run 3 406

H.24 DTM created using Settings 2 Run 3 406

H.25 Classification created using Settings 3 Run 1 408

H.26 Terrain points created using Settings 3 Run 1 408

H.27 DTM created using Settings 3 Run 2 408

H.28 Classification created using Settings 3 Run 2 408

H.29 Terrain points created using Settings 3 Run 2 408

H.30 DTM created using Settings 3 Run 2 409

H.31 Classification created using Settings 3 Run 3 409

H.32 Terrain points created using Settings 3 Run 3 409

H.33 DTM created using Settings 3 Run 3 409

H.34 Classification created using Settings 3 Run 3 410

H.35 Terrain points created using Settings 3 Run 3 410

H.36 DTM created using Settings 3 Run 3 410

H.37 Classification created using Settings 4 Run 1 412

H.38 Terrain points created using Settings 4 Run 1 412

H.39 DTM created using Settings 4 Run 2 412

H.40 Classification created using Settings 4 Run 2 412

H.41 Terrain points created using Settings 4 Run 2 412

H.42 DTM created using Settings 4 Run 2 413

H.43 Classification created using Settings 4 Run 3 413

H.44 Terrain points created using Settings 4 Run 3 413

H.45 DTM created using Settings 4 Run 3 413

H.46 Classification created using Settings 4 Run 3 414

H.47 Terrain points created using Settings 4 Run 3 414

H.48 DTM created using Settings 4 Run 3 414

List of Tables

4.1 Parameters for capture of EA laser data for pilot study 118

5.1 Statistics of the Differences between the Airborne and Terrestrial Laser Data . . . 168

6.1 Results Table for Split Sample Validation of Each Method Using 1m Resolution Data185

6.2 Results Table for Split Sample Validation of Each Method at Different Grid Spacings186

6.3 Descriptive Statistics for the Spline Surface With and Without a Buffer 189

6.4 Pilot Study Results of Cross Validation Analysis 189

6.5 Pilot Study Results for Jackknife Technique 192

7.1 Results showing the errors identified using the split-sample test for the main investigation. Surfaces generated using 95% of the data 201

7.2 Results showing the errors identified using the split-sample test for the main investigation. Surfaces generated using 50% of the data 201

7.3 Results showing the errors identified using the split-sample test for the main investigation. Surfaces generated using 25% of the data 201

7.4 Contingency Table Showing Accuracy of Aerial Photograph Classification 214

7.5 Validation with terrestrial laser dataset 214

8.1 Parameter settings for Image Division 228

8.2 Parameter Settings for default and adapted ATM algorithm 247

8.3 Contingency Table for the adapted IRI filtering method for the Bicubic 1m dataset 255

8.4 Contingency Table for adapted IRI method for the Nearest Neighbour 1m dataset 256

8.5 Results for Reliability Test for Classification Accuracy Assessment 258

8.6 Overall accuracy results table for IRI method 260

8.7 Overall accuracy results table for ATM method 261

8.8 Overall accuracy results table for Geometric method 262

8.9 Descriptive Statistics for Image Differencing (Part I) 282

8.10 Descriptive Statistics for Image Differencing (Part II) 283

8.11 GPS Comparison Results: Geometric Algorithm 285

8.12 GPS Comparison Results: ATM Algorithm 285

8.13 GPS Comparison Results: IRI Algorithm 286

8.14 Summary matrix of hypotheses results 289

8.15 Z-test Results for Comparison of Different Settings for the Three Algorithms . . . 290

8.16 Hypothesis Results of the Sensitivity Study 291

8.17 $Z - test$ and hypotheses results to show the effect of data structure on classification accuracy of filtering algorithms 293

8.18 $Z - test$ and hypotheses results for default IRI algorithm 294

8.19 $Z - test$ and hypotheses results for adapted IRI algorithm 294

8.20 $Z - test$ and hypotheses results for default ATM algorithm 295

8.21 $Z - test$ and hypotheses results for adapted ATM algorithm 295

8.22 $Z - test$ and hypotheses results for default Geometric algorithm 295

8.23 Z-test and hypotheses results for adapted Geometric algorithm 296

8.24 Comparison of ATM default results 297

9.1 Contingency Table Showing Segmentation Accuracy for Linear 0.5m Dataset . . . 307

9.2 Producer's Accuracy Results for Segmentation Using the Geometric Algorithm with Default Settings 309

9.3 Producer's Accuracy Results for Segmentation Using the Geometric Algorithm with Adapted Settings 310

9.4 Producer's Accuracy Results for Segmentation Using the ATM Algorithm with Default Settings 310

9.5 Producer's Accuracy Results for Segmentation Using the ATM Algorithm with Adapted Settings 311

9.6 Results for assessing height accuracy of segmentation procedure 321

9.7 Results of comparison of raw data segmentation with gridded data segmentation for the ATM method with default settings 324

9.8 Results of comparison of raw data segmentation with gridded data segmentation for the ATM method with adapted settings 325

J.1 Producer's Accuracy Results for Geometric Algorithm 418

J.2 Producer's Accuracy Results for ATM Algorithm 419

J.3 Producer's Accuracy Results for IRI Algorithm 420

M.1 Results for default IRI algorithm 435

M.2 Results for adapted IRI algorithm 436

M.3 Results for adapted ATM algorithm 436

M.4 Results for default Geometric algorithm 436

M.5 Results for adapted Geometric algorithm 437

N.1 $Z-test$ and hypotheses results for Geometric algorithm 439

N.2 $Z-test$ and hypotheses results for ATM algorithm 440

O.1 Results for Geometric algorithm using Default Settings 442

O.2 Results for Geometric algorithm using Adapted Settings 443

O.3 Results for ATM algorithm using Default Settings 444

O.4 Results for ATM algorithm using Default Settings 445

P.1 Comparison of Above Ground Classification Accuracies for Geometric Algorithm with Default Settings 446

P.2 Comparison of Above Ground Classification Accuracies for Geometric Algorithm with Adapted Settings 447

P.3 Comparison of Above Ground Classification Accuracies for ATM Algorithm with Default Settings 447

P.4 Comparison of Above Ground Classification Accuracies for ATM Algorithm with Adapted Settings 447

Acknowledgements

Huge thanks to my amazing principal supervisors, Professor Paul Longley (UCL) and Dr David Holland (Ordnance Survey). You have both been fantastic support, very constructive, and extremely patient! I am honoured to have had you both as mentors, and as friends. To David, special thanks for all of the discussions regarding research focus and for the technical help. You are a fantastic teacher, and incredibly understanding. I cannot thank you enough. To Paul, thank you so much for providing research direction and for the drive to make the Ph.D. happen in the first instance.

I would like to express my gratitude also to Professor Mike Batty, my second UCL supervisor, who provided much needed advice and support throughout the whole journey.

I am extremely grateful to Ordnance Survey who sponsored the Ph.D., especially to those integral in setting it up: Sallie White (née Payne), Dr Chris Parker, Dr Nick Land, and Ed Parsons. Thank you also to Dr Doreen Boyd who, over the last year at Ordnance Survey, has been extremely supportive and offered some great advice. Namely, hurry up and get it done!

Thanks to all members of the Research and Innovation Department who have put up with me being so stressed. Special thanks to Dr Isabel Sargent for help with MatLab and LaTeX (I have learned so much from you, thank you), to Dr Paul Cruddace for GPS advice and helping set up terrestrial GPS and laser scanning, and to Jon Horgan for creating the photogrammetric DTM.

Thanks to the Environment Agency, and to Infoterra for supplying airborne laser data used in this research.

And finally I would like to express my gratitude to:

My parents, Helen and Martin, and to my brother Stephen for their unfailing faith and support, and for always, always being there.

My friends for listening to my woes, time after time. Special thanks to Rach, Giles, Kat, Lee, Anne, Tony, and last but certainly not least, Tobes.

And especially to Gary - for being brave throughout the whole thing. Words cannot begin to describe how wonderfully supportive you've been throughout everything. Thank you.

This thesis is dedicated to my Grandad, Antony Urwin (1926 - 1991)

Chapter 1

Introduction

> Error-laden data, used without consideration of their intrinsic uncertainty, are highly likely to lead to information of dubious value. *Zhang and Goodchild* (2002) *pg. 3*

Error and uncertainty is almost inevitable in any representation of reality, and elevation models derived from Airborne Laser Scanning (ALS) systems are no exception. Whilst the sources of error during laser scanning data capture have been intensively investigated in the past (e.g. Huising and Pereira 1998b, Baltsavias 1999), insufficient attention has been given to the many error sources in the post-processing of laser data by end-users, and by implication, to the quality of the height models which are produced.

The total error in an ALS model is a function of the errors introduced by pre-flight system calibration (including georeferencing errors), data capture, preliminary post-processing, and end-user post-processing. Without quantification of the errors associated with end-user post-processing a full error budget for laser scanner derived models cannot currently be calculated. The purpose of this thesis is to estimate the error incurred during end-user post-processing. This focus is shown diagrammati-

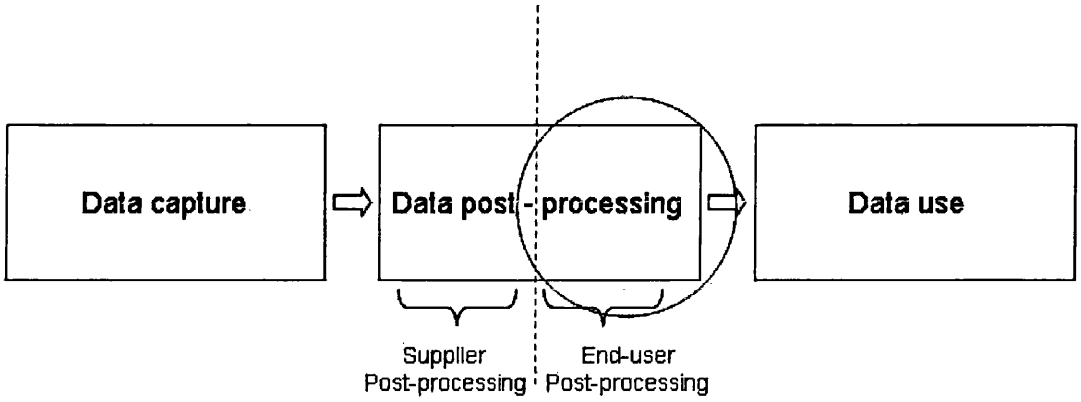

Figure 1.1: Workflow diagram showing the processing of laser data from capture to user application. Red circle indicates principal focus for thesis research

cally in Figure 1.1, in which a simplified workflow for ALS processing is depicted. It can be seen that ALS data processing is split between that conducted by the supplier, and that done by the data user. Emphasis in this thesis is on the latter.

The findings from this investigation will allow future users of ALS data to combine the information about post-processing errors with that currently known about the errors in data capture and in preliminary post-processing. This will enable them to both manage and minimise error, but will also allow them to estimate the quality of the models they produce. In this thesis the term 'user' is taken to mean any person or company who needs to process point cloud laser data. This may range from an individual using the data for a specific project (for example to create height models for modelling noise in the urban environment) to a large organisation wishing to process the laser data as part of a production flowline.

The specific focus of this research is on understanding ALS post-processing errors in urban areas. Information about uncertainty and error is particularly important for models of urban elevation, as here errors may be numerous given the complexity of this environment. Many authors researching the creation of models from ALS data have noted difficulties in these areas, underlining the necessity for research in

this field. The analysis of ALS models has hitherto been the preserve of the remote sensing community. Further understanding of these models may be attained through collaboration with research fields in which urban applications and urban processes are also investigated. The work of this thesis opens the debate on uncertainty modelling to a far wider range of relevant applications.

This chapter introduces the research problems which have given rise to this thesis, the stages of user post-processing are introduced, and the aims of the research presented.

1.1 Background and the Choice of Airborne Laser Scanning

The need for high resolution, accurate elevation models of the urban environment is growing, fuelled largely by demands from the insurance, telecommunications, planning, construction, architectural, and mobile technology sectors of the market. Each of these sectors requires highly accurate and precise models (in the range of centimetres) which can be used for detailed analysis and decision making. Current demands are for continuous models of terrain, surface and objects, and for discrete vector models of buildings, vegetation, and street furniture. In many applications the common requirements for both types of models are that they need to be produced quickly, be competitively priced, and include an indication of model accuracy. Airborne laser scanning is a remote sensing data capture technique which can potentially meet many of these demands. It is an active remote sensing technique in which a pulse of light is emitted. When the emitted pulse coincides with an object some light will usually be reflected from this. The reflected pulse is recorded by the sensor, and the time between pulse emission and the recording of the return pulse is measured. This is combined with other information from Global Positioning System (GPS) and Inertial Navigation System (INS) units and transformed to a distance measurement which gives the height of the surface at the point of coincidence with the laser beam. The raw ALS data are supplied as (x, y, z) points in 3D space, in

what is frequently termed a point cloud. Chapter 2 describes the fundamentals of laser scanning in more detail.

The commercial use of airborne laser scanning (ALS) for the rapid creation of such models has gained wider acceptance in the last few years because of its suitability for capturing information, particularly in complex urban areas. Advances in laser scanning technology have led to the development of more reliable and accurate systems (Sithole and Vosselman 2003), and the widening availability of data in many countries. In addition, laser scanning offers competitive data capture accuracies in the order of 0.5 - 1.0m in x and y, and 10 - 15cm in the z dimension (based on a flying height of 1000m above ground, Maas 2000) making it an attractive source of data for urban remote sensing and height modelling, and currently the most likely to provide a solution for the demands of the market place. In terms of the speed of production, there has been a considerable research effort investigating the automation of the creation of elevation models and the extraction of objects. Indeed, the issue of the automation of modelling from ALS data has been the primary focus of research in this area for many years (Vosselman 1999; Sithole and Vosselman 2003; Dash et al. 2004). Some of these processes have already begun to enter the commercial arena, including the automatic creation of terrain models from ALS data in some software packages (for example SCOP++ (Inpho) and TerraModeller (TerraSolid)).

In terms of pricing, it is widely acknowledged that laser scanning systems offer a cost effective means of acquiring digital elevation data (Flood and Gutelius 1997). Several practical investigations have been conducted to quantify this cost efficiency. Barnsley, Steel, and Barr (2003) quote an example from a report by the Environment Agency of England and Wales (EA 1998) in which the capture costs of elevation of both laser scanning and photogrammetry were compared. It was found that the cost of capturing elevation using photogrammetry was five times greater than capturing elevation for the same area using laser scanning. Other investigations offering similar conclusions include Lohr (1998) and Baltsavias (1999), the latter of whom concluded that in many instances laser scanning not only competes with photogrammetry as

a data capture technique for height, but in some cases should replace it.

The third requirement for ALS elevation models noted above was the provision of an indication of model accuracy. A large amount of research has attempted to model the errors associated with data capture (Huising and Pereira 1998b), including the modelling of the errors associated with each of the system components (see also Chapter 2). As such, the accuracy of raw ALS data is reasonably well understood, and some (global) measure of this is often provided with the raw data by suppliers. Many authors have also examined the errors incorporated into the datasets by preliminary post-processing techniques and have published some estimates for errors produced by these processes (e.g. Crombaghs et al. 2000; Maas 2000; Schenk et al. 2001). However, the same emphasis has not been placed on the understanding and modelling of errors in the end-user post-processing of raw ALS data, and there has recently been a call for the supply of such quality information (Sithole 2003). The issue of accuracy is particularly important given that ALS models, which are the product of end-user post-processing, are used for such a variety of applications including the surveying of coastal erosion (Palm and Lohr 2000), urban modelling (Palmer and Shan 2001; Haala and Brenner 1997) disaster management (Steinle and Voegtle 2001) the mapping of linear corridors such as electrical transmission lines (Wehr and Lohr 1999), erosion monitoring surveys, glacial surveys, volume control (for open pit mining), and flood prediction and management (Cobby et al. 2001). Emerging techniques using full waveform analysis of the return laser pulse to investigate details of objects, such as the structure of forest canopy, are also receiving greater attention as the technology gains further acceptance.

Other major applications include policy-making, quantitative analysis, and visualisation. The quality of such analysis is dependent on the accuracy of the models, as any errors in the underlying ALS model will be propagated through to these applications. For a user to assess the quality of the laser data and the models produced from these, some quantification of the errors incorporated at each stage of this workflow is required, and yet many users are unaware of the quantity and location

of errors which may be introduced during post-processing and modelling (Wechsler 2003). The principal post-processing stages for ALS data are introduced below.

1.2 Post-Processing of Airborne Laser Scanning Data

Post-processing of ALS data entails the creation of entire datasets (including the matching of strips of laser data, Maas 2000; Behan 2000), the modelling of systematic, random and human errors in data capture (Huising and Pereira 1998b), filtering of laser points, segmentation of data, feature detection, and 3D reconstruction. These processes are shown in Figure 1.2 along with the products derived from each. The post-processing tasks shown in Figure 1.2 can be divided into two distinct groups: firstly, the immediate post-processing of data following capture in order to render data suitable for supply. This includes the transformation of information into common coordinate systems, the computation of distance values for each point, and the matching of laser strips (see Chapter 2 for more information about these processes). Secondly, end-user post-processing is conducted by data-users following supply. It is these later stages of processing which are investigated in this thesis, and which are described in the following sections.

1.2.1 Data Structuring

An ALS data user must first decide upon the data structure in which they wish their data to be held. Different data structures require different approaches to modelling. These different approaches may create varying amounts of error during post-processing. It is important for users to understand these differences if they are to use those processing techniques which are most appropriate to their requirements.

Post-processing can be performed either on the raw points, which may be discrete points or may be connected together as a Triangular Irregular Network (TIN) , or

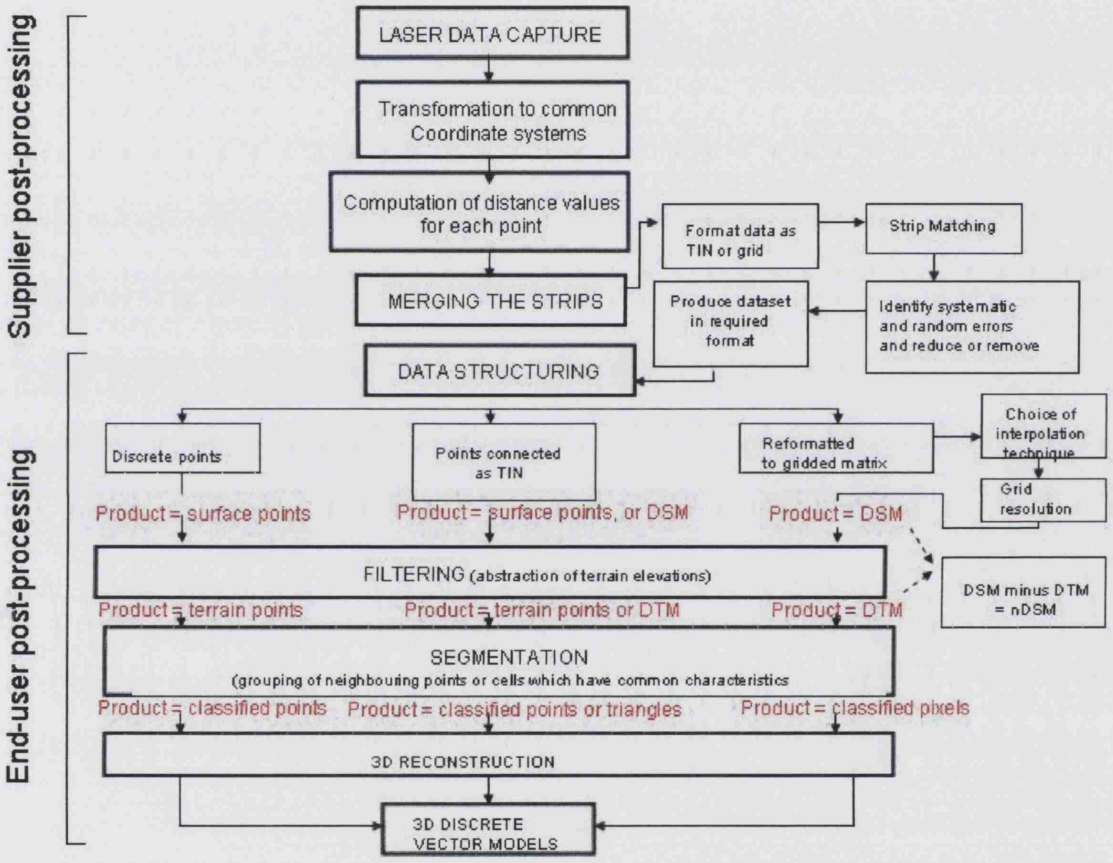

Figure 1.2: Flow diagram showing stages in the post-processing of Airborne Laser Scanner data

they may be re-structured onto a gridded matrix. Where they are connected in a TIN the values of the raw elevation data are generally not changed. However, errors may be introduced in the interpolation of elevation in the intervening non-sampled locations. The characteristics of these errors will be dependent upon the interpolation algorithm used. In most cases this is a linear interpolation (Kumler 1994), however alternative methods have been used in the past (these are reviewed in McCullagh 1988).

The grid data model requires that the original data are restructured onto a regular grid using some form of interpolation procedure. The advantage of this approach is that the data can be modelled and analysed using standard image processing techniques, and that the structure is acceptable in most commercially available software. The gridded matrix as a structure for elevation datasets has also been widely accepted as a standard in elevation modelling (Burrough 1986; Carter 1988; Weibel and Heller 1991), and particularly for analysis in geographical information systems (Weibel and Heller 1990, 1991; Wood 1996). The grid structure is also required for a number of specific ALS procedures, including the measuring of planimetric and vertical shifts between overlapping strips of laser data (Behan 2000). Despite the small separation distances of the laser points, the interpolation procedure will inevitably introduce error into the dataset which is a major disadvantage of this approach. On the basis of this criticism a number of authors have chosen to avoid using this format in elevation modelling (e.g. Morgan and Habib 2002; Pfeifer et al. 2001a; Schinckler and Thorpe 2001), however a large number of proponents of the grid approach remain (Brovelli et al. 2002; Elmqvist et al. 2001; Wack and Wimmer 2002) and its popularity for many applications continues. As part of the overall investigation, the importance of the choice and the method of data structuring is investigated. The implications of this for the subsequent stages of filtering and segmentation are also examined. The different approaches will introduce varying amounts of error into the derived models. This requires quantification if users are to decide which processing techniques are optimal for their requirements.

1.2.2 Data Filtering

Filtering is the abstraction of the points on the bare earth from an ALS point cloud (Sithole and Vosselman 2003), where the bare earth is defined as the terrain (topsoil or any thin covering, such as asphalt, pavements) without any above ground objects (such as buildings and vegetation). Such processing is required because the points, as supplied, are distributed both over the terrain surface and on objects above the bare terrain (such as trees, buildings, cars). Many applications require only a terrain surface for analysis. Datasets containing only bare earth points can be produced by filtering the raw points to retain only those which lie directly on the bare earth. As can be seen from Figure 1.2 the results of the filtering are used to create derived models such as the DSM, the DTM, the nDSM, or discrete point datasets of surface, terrain and objects. These products are used, in turn, in the creation of 3D vector reconstructions of objects within the urban landscape (such as buildings, and trees etc). Thus, the understanding of the characteristics of the error introduced during the filtering stage of modelling is crucial as this error not only influences the quality of the filtering, but will also be propagated to the derived products. It has been estimated that of all of these post-processing tasks the filtering of the laser data, and the quality control of this, pose the greatest challenges. (Flood 2001) has estimated that the filtering stage consumes some 60% to 80% of processing time.

1.2.3 Segmentation

Segmentation, or clustering as it is sometimes called, is the classification of neighbouring laser points, or pixels, having common characteristics which define them as being 'above ground'. This technique is used in the identification of buildings, trees and other surficial terrain objects, and directly follows the filtering stage in the processing of data. The processes of filtering and segmentation are inextricably linked as the points or cells rejected in the filtering process are those which are used in the segmentation process. Segmentation techniques are mandatory for the recon-

struction of objects, and for this reason most studies have focused on investigating procedures for urban areas.

In accordance with the processes defined above, there are a variety of elevation datasets that may be produced by end-user post-processing, and these require some definition. Where the heights in the model represent the terrain plus any objects on this (such as buildings, cars and vegetation) it is termed a Digital Surface Model (DSM). Where the heights in the model only represent the bare terrain it is called a Digital Terrain Model (DTM). Both the DSM and the DTM are forms of Digital Elevation Models (DEMs). DEM is a generic term used to describe a surface composed of regular or irregular point data, without distinction. The DEM is a representation of the elevation of the Earth's surface in digital form by a set of points which are arranged in a regular or irregular pattern. Until recently it was generally considered a DEM could only refer to arrays of regular points, however with the proliferation of the use of laser data (which is of course irregular) the term has been widened to include irregular datasets, such as in Buckley and Mitchell (2004). Therefore, DEMs are discrete representations of the Earth's continuous surface. Derived elevation models can be used to produce continuous height models (normalised Digital Surface Model, nDSM) in which unknown heights are interpolated from known points , or they can be used to create 3D vector reconstructions of objects. A number of research areas have started using the terms Digital Elevation Data which is far more sensible, but using this is likely to cause more confusion in the short term.

Previous methods for modelling errors and the approach taken in this research are discussed below.

1.3 Modelling the Errors in Urban ALS Models

1.3.1 Definition of Terms

This thesis uses a number of terms to describe data quality, these include: uncertainty, error, and accuracy. These have been defined previously in Wood and Fisher (1993), Wood (1994), Foody and Atkinson (2002), and Zhang and Goodchild (2002). Whilst the definitions proposed by these authors largely concur there are some subtle differences. For remote sensing research those presented by Foody and Atkinson (2002) appear to be the most suitable.

Data quality refers to the suitability of data to be used for a particular application. It is often defined in terms of error and accuracy. Foody and Atkinson (2002), and other authors, define error as the difference between a true value and the prediction of this, and in accordance with this definition error relates to individual values.

$$\varepsilon_i = l_i - x_i \tag{1.1}$$

where:

ε_i = actual error of the observation l_i (some authors call this true error), l_i = observation of parameter at i, and x_i = true value of l_i. In other words, ε_i is the residual at i.

One conceptual problem with this definition of error which has been identified within the literature is that this definition assumes that the 'truth' is obtainable (Zhang and Goodchild 2002). However, any measurement of the real world is subject to a number of sources of error including systematic error, instrument mis-calibration, or human blunders. As such, measurements of the real world are simply representations or abstractions, and since it is rarely known, the true value is only of theoretical

significance (Heuvelink 1998). Error is thus defined here being the difference between the observed parameter and the estimate of the true value of the parameter.

$$\nu_i = l_i - \hat{x} \tag{1.2}$$

where:

ν_i = residual error of the observation l_i

$\hat{x}$ = estimate of true value l_i

It is worth noting here that this does not, of course, deal with the more philosophical issue of what reality is, and how we can begin to measure it. Such a discussion is beyond the remit of this thesis.

Accuracy is defined in terms of bias and precision (an accurate measurement must be unbiased and precise). Bias is an expectation of over- or under- prediction based on some statistical model (often mean error), whilst precision is the expectation of the spread of errors (Foody and Atkinson 2002). Accuracy is a global descriptor which is often based on a computed statistic of interest. In other words accuracy is an expectation of overall error (Foody and Atkinson 2002). In DEM research, accuracy is often used to describe the general characteristics of a surface, whilst error (which relates to individual values) is generally rarely estimated. Precision and bias are not focused upon in this research, these are more relevant to the modelling of errors in data capture.

In terms of this investigation, the error is estimated for each laser point. The calculated error per point is represented in terms of dot maps and isolines, these map the error. In the areas which are predicted to contain higher errors, the modelled surface is more difficult to estimate, and in this sense the model is more ambiguous. Maps to depict this ambiguity are used to show anticipated uncertainty in the model.

Mapping and visualising uncertainty has been an active area of research over the last few years, and studies such as Appleton et al. (2004) and Schneider (2004) have presented a number of techniques. The methodology used for error and uncertainty mapping as part of this investigation is detailed in Chapters 6, 7 and 10.

1.3.2 Methods for Modelling Error

To a large extent, modelling the error in ALS derived elevation models is no different from modelling the error in any other surface model. That said, the focus of previous research has posed some problems in terms of the transferability of existing methodologies. Estimates of error in elevation models have, in the past, generally been restricted to analysing the nature of the errors within bare earth surfaces in predominantly natural environments (e.g. Xie et al. 2003; MacEachren and Davidson (1987)). Very few studies have begun to consider the problem of understanding error in urban surface models, largely because of the inherent complexities of the urban surface with its frequent discontinuities, large variety of complex shapes and textures, and high rate of surface change.

Many previous investigations assessing the relative merits of different DEM methods have only looked at issues of global measures of error (Declercq 1996) such as root mean square error (RMSE) and error matrices. These measures have no spatial dimension (Wood and Fisher 1993), and yet the spatial pattern of errors can, in fact, reveal useful information about the phenomenon of interest as highlighted by Burrough and McDonnell (1998) in the quotation below:

> The usual view of errors and uncertainties is that they are bad. This is not necessarily so, however, because it can be very useful to know how errors and uncertainties occur, how they can be managed and possibly reduced, and how knowledge of errors and error propagation can be used to improve our understanding of patterns and spatial processes. Linking a

> good understanding of spatial uncertainties to numerical methods of modelling and interpolation can provide useful tools for optimising sampling (and thereby improving value for money) and for identifying the weak and strong parts of spatial analysis. A good understanding of errors and error propagation leads to active quality control.
>
> Burrough and McDonnell (1998) pg. 221

Maps showing the spatial distribution of errors may be analysed in terms of the patterns of over- or under-prediction, or by examining the patterning of the errors using indices of spatial autocorrelation (Fisher 1988), or the development of local indices of spatial association (e.g. Anselin et al. 2003). A better understanding of error and uncertainty is fundamental to the future of ALS modelling in both commercial and academic fields. This error information must be presented in such a form as to be useful to ALS data users so that it may be incorporated within the decision making process and eventually may be used to reduce errors in output products (Zhang and Goodchild 2002). Making users aware of potential errors or the general 'quality' of a representation constitutes an important decision support tool.

1.4 Motivation For This Research

1.4.1 Problem

The necessity for this research in the ALS community is driven by the fact that the accuracy of elevation models will directly affect the quality of any analysis based on the modelled height values. ALS derived elevation models are used in many analytical operations such as slope and aspect calculations, and processes such as image segmentation and filtering in the creation of Digital Terrain Models (DTM). When combined in applications these operations can be highly sensitive to

the quality of the elevation data used (Wood 1994; Florinsky 2002). However, many end users of height models in geographical information systems (GIS) are unaware of the issues surrounding the quality of the underlying height data, partly because of the lack of diagnostic tools for dealing with quality information in many commercial software packages (Desmet 1997). Even where users are aware of quality issues with the DEM, there is often a lack of methods available to accommodate them in subsequent spatial analysis (Goodchild and Gopal 1989). The sum consequence is a rather limited understanding of the uncertainties inherent in applications using elevation models based on ALS data. It is worth noting here that users do not currently have this error information for height models created from other sources (such as GPS and photogrammetry), in this respect the lack of information for ALS models is not unique. It does, however, highlight the need for research in this area in general and for this reason many of the techniques presented are transferable. The emphasis placed on ALS data in this research is based on the current trend towards the increased usage of ALS data for deriving height models.

1.4.2 Drivers

This research was driven by both academic curiosity and commercial interests. The original impetus was largely academic, and arose from a recognition that despite the recent focus for urban elevation modelling on ALS the vast majority of this research was concerned with the derivation of new algorithms for feature extraction and elevation model creation. This focus had resulted in a dearth of research investigating the sources and operation of errors within, and the resultant quality of, ALS models. As shown above, this issue of uncertainty has become more commercially important as the demands from users for greater accuracy have increased, and the absence of research in this area perhaps presents a potential obstacle to the widespread uptake of laser scanning in the commercial arena.

The research presented within this thesis falls into the category of applied research.

The use of such an applied approach is relatively rare in the field of laser scanning research, which has tended to be dominated by pure research to date. Whilst the result of this emphasis on pure research is, on the one hand, a huge body of literature and knowledge, on the other hand the lack of applied geographical approaches has meant that this knowledge has rarely been linked to real world problems. In recent years, there has been a change in emphasis in some research areas towards extolling the virtues of laser scanning for applications such as forestry inventories, and whilst these represent a move towards an applied approach they often fall short of providing useful information for users at a decision making level (as implied by Wechsler 2003). Applied research (Briggs and Butterworth 1981) continues to be rare in the laser scanning community. This is strange given the undeniable applicability of the technique. The lack of applied research is likely to be the result of the current focus of research on creating algorithms which can achieve a higher level of automation for ALS processing. This thesis represents an attempt to demonstrate how an applied research approach could, in fact, enhance the uptake of ALS data for urban height modelling, which would in turn stimulate the need for more pure research. The future success of the laser scanning academic community relies on a more even balance between pure and applied approaches.

1.5 Aim of the Thesis

The thesis underpinning this research is that different post-processing techniques for ALS data will produce varying amounts of error, both in magnitude and in location. The implication is that this error will subsequently be transferred to any analysis made on the results of the post-processing. For this reason, ALS data users should aim to reduce error within models by choosing the most appropriate processing algorithm(s) for their modelling. In order to do this, more information is needed regarding the sources and the operation of errors in ALS datasets. The research presented here is concerned with the quantification of the errors within

post-processing. The techniques of data structuring, filtering and segmentation are analysed to assess the magnitude and location of errors introduced by each. The impact of using different algorithms for each of the stages is examined. It is hoped that answers to these issues will provide information and a framework for future ALS data users to facilitate decision making in the design of their post-processing workflow.

The results of the work quantifying error will be of use to those ALS data users wishing to manage error in the models they produce. In particular, the investigation is designed so that the results provide answers to the following unresolved research questions in laser scanning research:

- How does the choice of structure for ALS data affect the magnitude and the spatial patterning of errors in derived elevation models?
- What are the characteristics of the error introduced by the interpolation of point cloud data onto a regular grid? How does the choice of interpolation method and grid resolution affect the characteristics of the error introduced?
- What does the spatial pattern of errors reveal?
- Does the choice of filtering algorithm make a significant difference to the accuracy of the resultant datasets?
- How sensitive are the processes to changes in parameter settings?
- Does the spatial resolution of the data affect the characteristics of the error in the resultant datasets?
- Does the accuracy of segmentation datasets vary according to algorithm used?
- Is there a significant difference between the processing of raw laser data and gridded data
- How might the uncertainty in ALS models be measured, visualised and communicated to the user?

Some of the modelling methods compared in this thesis are new algorithms, written as part of this research, whilst others are based on commercially available methods which have been adapted for use in this study.

1.6 Thesis Outline

The thesis has two main parts, with the first presenting background information and the motivation for research, and the second presenting the research and conclusions.

The first part of the thesis comprises Chapters 1-4, in which Chapter 2 presents the technical aspects of airborne laser scanning, including descriptions of the system components and the sources of errors within each of these. The equations for calculating laser point coordinates and values, and the full error budget equation for modelling errors in ALS data are detailed. This equation is presented for the first time in this thesis, it includes parameters for errors in post-processing in addition to those which have already been quantified for pre-processing. The quantification of each of the post-processing parameters is the focus of Chapters 6-9.

Chapter 3 presents the research context in which this research has arisen. The three principal areas of literature which have influenced this research are identified and reviewed. Firstly, previous work investigating different representations of urban form is discussed. This sets the research in a much wider academic context. Here parallels are drawn between this research, which focuses on the modelling of urban elevation from ALS data, and surface representation in other disciplines. Secondly, previous elevation modelling techniques are reviewed and thirdly, urban remote sensing is discussed, with focus on applications and problems of various techniques.

Chapter 4 describes the data used in the investigation and presents the methodologies used for quantifying and analysing error. The mathematical theory of the spatial interpolation methods used within this thesis is introduced, as is the theory

of the validation techniques (such as cross validation and jackknifing) as these have not been widely used in laser scanning research previously. The final section of this chapter introduces the data investigated within the research, and describes the supplier post-processing techniques used in the creation of the raw dataset.

The second part of the thesis consists of Chapters 5 - 11 which present the results and findings of the three main investigations. In Chapter 5 the errors created during data capture are quantified, and it is shown how these results will be used in the later calculations of post-processing error.

Chapter 6 presents the results from the pilot study which was used to define the methodology employed in Chapter 7. Chapter 7 details the results of the main experiment assessing the impact of data structuring on the errors introduced. The sensitivity of the data restructuring methodologies to changes in parameters (including interpolation method and spatial resolution) is identified, and the spatial pattern of errors analysed.

Chapter 8 presents the results of an investigation comparing three filtering methodologies: two point based algorithms, and one grid based. The grid method algorithm is original and was written for this investigation, the point methodologies used are published methods which have been adapted for this investigation. The filtering methodologies are described, and the errors in each resulting model quantified and mapped.

In a similar way, Chapter 9 presents the results of the analysis of those errors introduced by segmentation.

Chapter 10 consists of an evaluation of the methods and results produced in the previous three chapters and offers an overview of the different methods and the errors produced. Final results of all comparisons are presented and conclusions drawn. The visualisation and communication of error information is briefly investigated.

Chapter 11 summarises the conclusions of the thesis, and discusses the contribution of the research and the transferability of the results to other areas of investigation. The continuing research themes which have arisen from the work are identified and provisionally assessed.

1.7 Chapter Summary

This chapter has illustrated how ALS data can be used to create a variety of elevation models which represent the heights of the upper physical surface, the terrain, and the heights of identified objects within the urban landscape. These elevation models are created using a variety of procedures, any of which may introduce error into the model. It is shown that the nature of this error remains largely unknown and that this has, in the past, created uncertainty for users of ALS data.

Chapter 2

The Theory of Airborne Laser Scanning

2.1 Chapter Introduction

The purpose of this Chapter is to introduce the theory of airborne laser scanning. This includes a description of the system components and the process of data capture. The characteristics of the raw data are also highlighted, and preliminary processing techniques, and potential error sources, described. Current and future applications of ALS data are discussed.

2.2 Laser Theory

Laser is an acronym which stands for light amplification by the stimulated emission of radiation. This describes very succinctly how a laser works. Although there are a wide variety of types of laser, they all have certain essential features and work on the principle of exciting atoms. When energy, in the form of heat, light or electricity, is applied to an atom it leaves what is called ground-state energy and enters an excited

state. In this excited state, the electrons orbiting the nucleus of the atom change their orbiting pattern and enter higher energy orbits. This change in orbit is less stable and eventually results in the electrons wanting to return to ground state. In order to do this, the atom releases energy in the form of photons, which are particles of light. In a laser, the way in which atoms release photons is controlled so that the light released is monochromatic (it has one specific wavelength), it is coherent (all photons move in the same direction), and it is directional (the light is concentrated into a very tight beam). All photons have a certain wavelength which is dependent on the energy difference between the excited and ground state. If a photon should collide with an atom which has an electron in an excited state, then the first photon can induce atomic emission. In other words, the collision of a photon with the atom causes another photon to be released from the excited atom. This photon has the same frequency and direction as the original photon. In a laser, these collisions are typically contained with a cylindrical tube. At each end of the tube is a mirror, at one end is a fully reflecting mirror, whilst at the other end the mirror is only partially reflecting (Jelalian 1992) meaning that it reflects some light but also lets some light through. An energy source is applied to the tube containing atoms, the electrons in the atoms become excited and some release photons. The photons travel within the tube and some are reflected off the mirrors and travel back into the cylinder, colliding with other excited atoms and stimulating the release of other photons of the same wavelength and direction. Upon collision with the partially reflecting mirror at the end of the cylinder, some of the photons escape. The light which escapes is laser light.

Laser technology is a rapidly expanding field of research with lasers being used in a wide range of applications including microsurgery using optoelectronics, defence technology, computer science, communications technology and remote sensing (Measures 1992). Its potential for remote sensing was first explored by Fiocco and Smullin (1963) who recorded laser echoes from the upper regions of the atmosphere and by Ligda (1963) who studied the troposphere. Since these pioneering investigations laser scanning has been used in a variety of studies which have exploited the ability

of laser systems to provide spectral information of distant targets (Measures 1992).

Lasers may operate at wavelengths in the ultraviolet, visible, near-, mid-, or far-infrared regions of the electromagnetic spectrum. They are used to transmit radiation to a target via an output optic (a telescope). The scattered target radiation can be detected and processed in a laser radar receiver to remotely determine the target characteristics (Jelalian 1992). Laser radars are effectively a direct extension of conventional radar techniques, the difference being that they operate at much shorter wavelengths than conventional radar which uses microwaves. Because they operate at shorter wavelengths, laser radars are capable of higher accuracy and more precise resolution than microwave radars, however they are more affected by atmospheric conditions and so tend to be used at shorter ranges in the lower atmosphere than microwave radar.

Laser radars tend to be called either LiDAR (which stands for both Light Detection and Ranging or sometimes Laser Induced Detection and Ranging) or LaDAR (Laser Detection and Ranging). The terms LiDAR and LADAR have tended to be used interchangeably in the literature to refer to both the technique and to the system itself, and this has led to some mis-understanding of the terms within the literature. To further confuse matters, the technique of laser scanning is also referred to as both laser ranging and laser altimetry.

2.3 Laser Scanning for Remote Sensing of the Earth

In the 1970s laser scanning began to be used for monitoring the solid Earth itself, as well as its atmospheric constituents, from mobile platforms including aircraft and helicopters. The first targets for observation were surface waves and coastal bathymetry (for example, Bunkin et al. 1983). These studies further demonstrated the potential of lasers for environmental remote sensing, and the flexibility of the technique which could be used from ground, air, ships or satellites at a range of

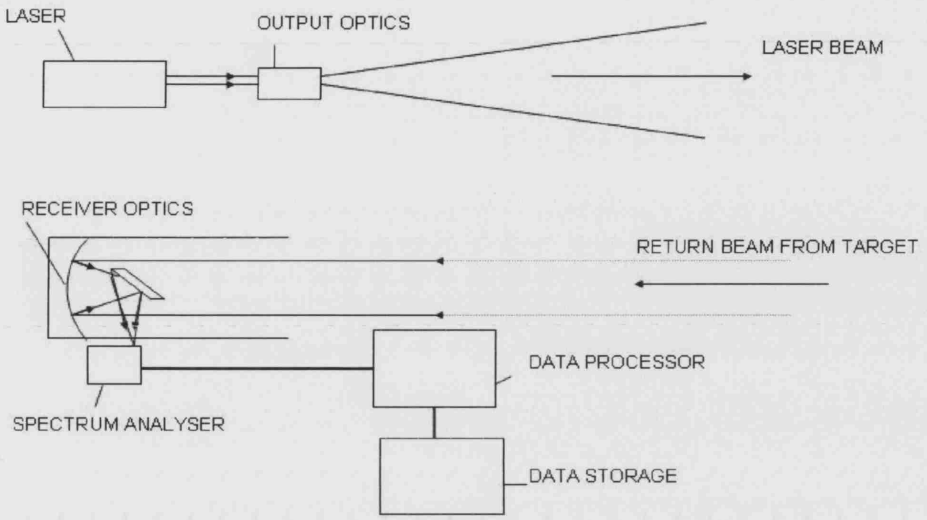

Figure 2.1: Elements of laser sensor (after Measures 1992)

both temporal and spatial scales. Lasers used for such remote sensing applications have to be high powered, with narrow bandwidth, short duration and display a low degree of diversion (Measures 1992).

Lasers are either continuous wave or pulsed, those used for commercial airborne scanning now are all pulsed and for airborne terrain mapping all use the near infrared (NIR) with a high repetition rate. The technique of Q-switching, which is attributed to McClung and Hellwarth (1962), made possible the generation of short, individual pulses of light which are used in airborne lasers today.

Figure 2.1 shows the set up of a typical laser scanner. The laser produces a pulse of coherent optical light which is directed through the output optic towards the target of interest. Radiation from the return pulse is recorded by the receiver optic system. The return pulse then passes through the spectrum analyser (to select the observation wavelength only and so discriminate against background radiation at other wavelengths), and finally to the photodetector. The signal from the photodetector is then analysed digitally.

The time between the emission of a single light pulse and the arrival of the return

signal can be directly related to the range if the velocity of light is taken as a constant. This is the fundamental principle underpinning airborne laser scanning theory. The system records the time taken for the emitted beam to return and be recorded by the receiver optics. This time measurement can then be converted into a distance, or range, measurement. The equations for this calculation are shown below.

$$2R = c_n(t_R - t_0) = c_n t_m \tag{2.1}$$

$$R = \frac{c_n t_m}{2}$$

$$\Delta R = \frac{(t_m \Delta c_n + c_n \Delta t_m)}{2}$$

Source: Lindenberger et al. (2003).

where: R is the range, c_n is the speed of light in air, t_m is the time elapsed between laser emission and detected return signal, $t_m \Delta c_n$ is the error in the speed of light calculation, $c_n \Delta t_m$ is the contribution from error in the timing within the system, and ΔR is the error in the range calculation

Not only can the range be measured, but the height of the surface of the Earth at specific locations can be calculated if the position of the mobile platform (in this instance the aircraft) can be ascertained. In current airborne systems this is achieved using two pieces of equipment - an Inertial Navigation System (INS) and a Global Positioning System (GPS). These, and other, system components used in airborne laser scanning are described further in the following section.

2.4 Topographic Airborne Laser Scanning

2.4.1 System Components

Airborne laser scanning can be conducted from any airborne platform, the most common of which are fixed wing aircraft and helicopters. The components of a typical fixed wing airborne system are shown in Figure 2.2.

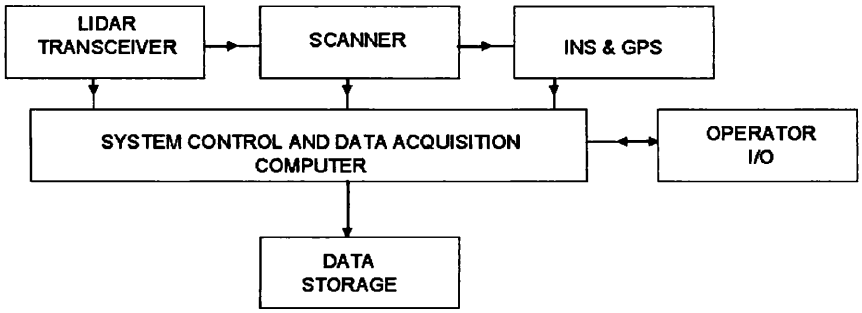

Figure 2.2: Airborne system functional block diagram

The lidar transceiver generates and emits a beam of light towards the scanner. The scanner then moves the laser beam across the flight track. There are three types of scanner which may be used in a system, these are: a rotating mirror, an oscillating mirror, and fibre optic scanners. The type of scanner affects the scan pattern of the resultant points. For example, a rotating scanner produces a regular squared scan pattern and a uniform measurement on the ground. Using this type of scanner some efficiency is compromised as a number of pulses are lost as a result of scanner geometry when the mirror rotates around one of its corners. An oscillating mirror produces a saw toothed scan pattern, in which the density of the laser points will vary across the scan line. Many of the commercially available fibre optic scanners produce a parallel strip scan pattern.

All commercially used ALS scanners, at present, require mechanical movement to reflect the laser beam repetitively across the flight line. The aircraft forward movement provides laser beam motion in the along-flight direction with the resultant rate of area surveyed (A)given by

$$A = 2vh \tan \Theta m^2 s^{-1} \quad (2.2)$$

Source: Lindenberger et al. (2003).

where: v = aircraft velocity (m/s)

h = aircraft altitude above ground (m)

Θ = optical scan-angle amplitude

The area of the ground covered by the scanning depends on the system field of view (FOV). The instantaneous field of view (IFOV) is determined by the receiver optics (focal length and detector size), while the FOV characterising the swath width is determined by the scanner mechanical capabilities - and is simply ± the maximum scan angle. Optech's ALTM is capable of scanning up to ± 30 degrees, however the most useful swath in practice is ± 25 degrees.

The position of the aircraft is determined by the GPS and INS units, together these are called the Position and Orientation System (POS). The GPS unit permits the accurate and reliable control of scanners and lasers in aircraft, and the INS (Inertial Navigation System) allows accurate aircraft position and orientation to be ascertained. The GPS subsystem typically consists of an antenna and a receiver, in addition to the ground based stationary receivers. The GPS antenna is situated on top of the aircraft to maintain satellite contact. The exact position of the antenna is recorded and used to calculate the position of the sensor. The difference between antenna and sensor is called GPS antenna eccentricity, and is measured after system installation and is taken into account either by INS data processing or by laser point processing. The airborne receiver provides the raw 'positional' data which consists of a time event marker, in addition to the GPS time and the raw position file both associated with that marker. The GPS unit is also used to time-stamp both

the target range and the ALS angle measurements with the result that the target position is referenced to GPS time. The accuracy of the whole GPS unit is typically 3-10cm (z).

The INS unit is used to determine the orientation of the aircraft and therefore of the sensor. The INS consists of three accelerometers which are mounted along three orthogonal axes, and three gyros used to define the reference frame for the accelerometer in order that acceleration may be measured in three dimensions. The INS unit records the following information: the positional vector in an earth fixed reference frame, rotational accelerations, the gravity vector, Earth rotation, and the transport rate.

The measurements taken by the GPS and the INS systems are in different coordinates, the GPS unit uses WGS84 and the INS uses the Inertial Reference Frame (IRF). These require transformation in order that they can be aligned in one common coordinate system called the Earth Centred Earth Fixed Frame (ECEF)(see also Equation 2.7 page 76).

Finally, the ALS airborne computer controls all of the ALS subsystems. It also integrates and synchronises the data from all of the subsystems for each recorded laser point. These data are compiled using the data acquisition system and are converted to distance measurements using Equation 2.2 and packed, formatted and stored on high-capacity disks. A schematic of airborne laser data capture is shown in Figure 2.3

Whilst all remote sensing lasers employ this configuration, there is a variety of specific lasers which are used for different purposes. These include: atmospheric laser radars, bathymetric systems, differential absorption lidar (DIAL), space lidar, terrestrial lidar, and airborne topographic lidar. This research is concerned with airborne topographic lidar.

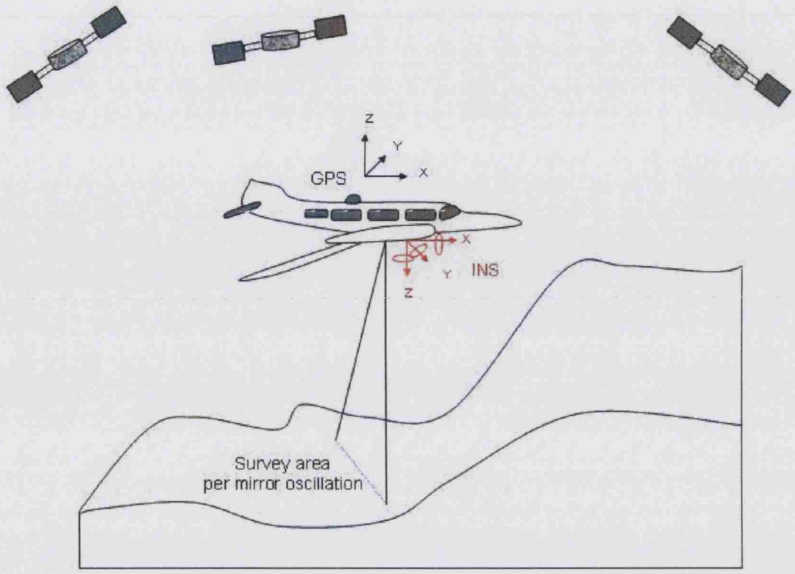

Figure 2.3: Airborne laser scanning

2.5 Characteristics of the Raw Data

The raw data stored on board the aircraft consist of three types of information for each point: height values, intensity values, and return attributes. The derivation of the height, or range, values from the time measurements was explained above. Figure 2.4 shows a map devised using a typical laser dataset with values coloured by recorded elevation. The intensity values and return attributes are discussed below.

2.5.1 Intensity of Recorded Pulse

The recorded return pulses vary in their intensity due to the differential spectral reflectance of different objects on the Earth's surface. An example of differential intensity is shown in Figure 2.5 in which the individual points have been coloured by intensity. In this figure different objects and land covers may be differentiated on the basis of differing intensity values. Differential intensity values can be used to enhance classification as has been shown by Song et al. (2002).

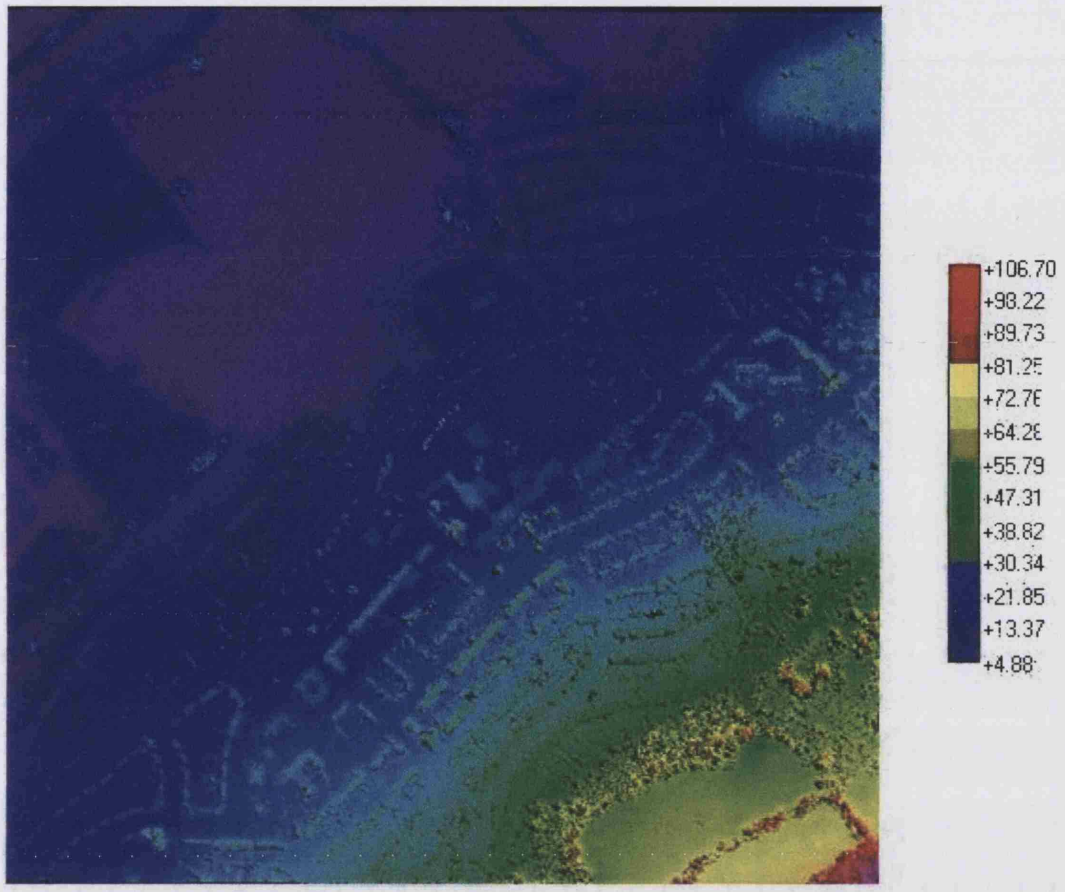

Figure 2.4: Laser data coloured by height (m), here displayed as an image

Traditional spectral reflectance curves for target types (as shown in Lillesand and Kieffer 2000 pg.17) show that water, for example, has low reflectance values. Laser scanning depends on the target surface being Lambertian-like (which is the case for most natural surfaces) rather than specular. As a result the maximum operational altitude decreases dramatically for steep, smooth, wet or black surfaces. This phenomenon can easily be seen in raw laser datasets where over water bodies values will generally be missing. In most applications of topographic laser scanning any recorded values over known water bodies should be ignored in the computation of any Digital Elevation Model. Varying reflectance of target objects is an important element of any flight planning, especially given that reflectivity can also change as a function of atmospheric conditions. For example, if there has been a period of sustained rainfall preceding flight over a target area it may be more productive to

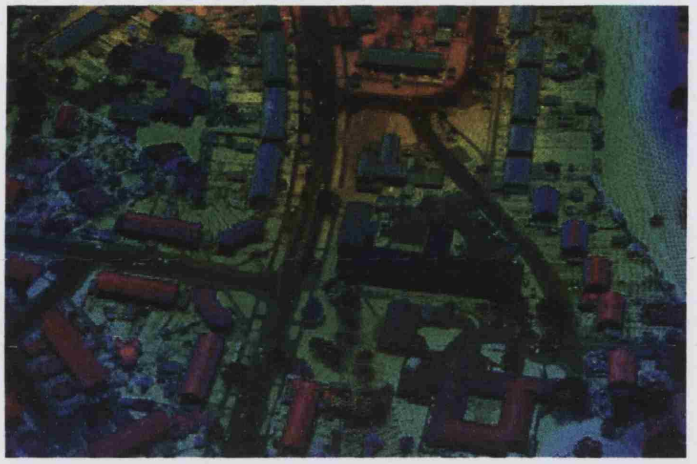

Figure 2.5: Laser data of subset of study area. Data coloured by intensity

fly at a lower altitude given the lower reflectivity of the damp surface. This highlights the intrinsic link between reflectivity, range and flying height. It should be noted here that if there is water lying on the surface then it is unlikely that any return pulse will be recorded as the surface will be more specular. The full equation for calculating the received signal power (P_s) from a transmitted laser pulse after reflectance from a target is given by:

$$P_s(R) = P_l \frac{P_t(R)}{\pi} \frac{A_r}{R^2} \eta_o \eta_a(R) \tag{2.3}$$

Source: Lindenberger et al. (2003).

where: P_s is the received signal power from the transmitted laser pulse after reflectance from target, P_l is the power of the laser pulse, P_t is the effective Lambertian reflectivity of the target, A_r is the effective collection area of the optical receiver, R is the slant range to the target from the sensor, η_o is the optical transmission efficiency of all optical components in the ALS, and η_a is the transmission efficiency of the atmosphere between sensor and target (at range R)

2.5.2 Return Attributes

The laser pulse produced by the transceiver will diverge upon emission from the sensor. The amount of this divergence (Θ_t) can be calculated and used to calculate the size of the laser footprint (S) which is a function of beam divergence

$$S \propto R \tag{2.4}$$

and

$$\lambda \simeq R\Theta_l \tag{2.5}$$

Source: Lindenberger et al. (2003).

where: S is the size of the laser footprint, R is the slant range from the sensor to the target,λ is the wavelength, and Θ_l is the laser beam divergence.

It can also be shown that:

$$\Theta_l = f(\lambda, D_l) \tag{2.6}$$

Source: Lindenberger et al. (2003).

where: D_l is the laser beam diameter at exit from ALS sensor

Figure 2.6 depicts a typical range measurement scenario.

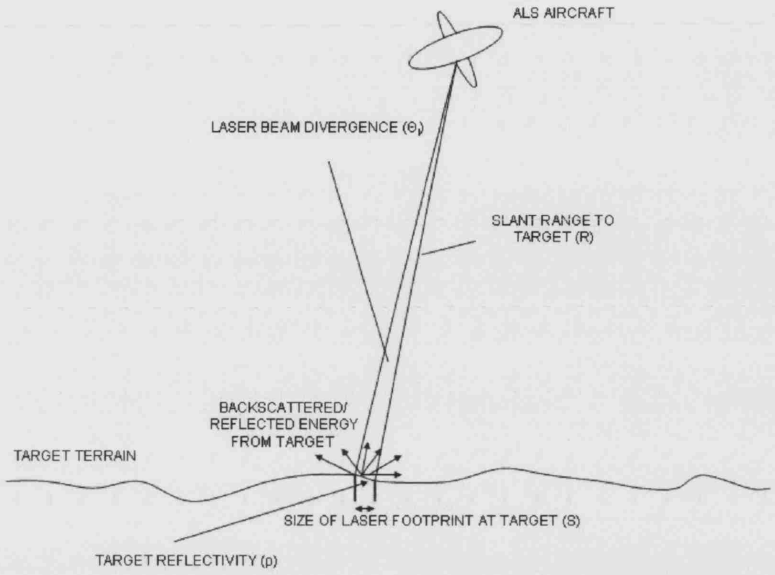

Figure 2.6: Typical Range Measurement Scenario (after Lindenberger et al 2003)

Each laser point therefore represents an average value over an ellipse. Where this laser pulse coincides with a flat surface a single reflectance value will be recorded indicating that the whole of the footprint hit a solid surface. In other situations the laser pulse may be split. This occurs when the laser ellipse covers more than one surface feature, such as where the target is smaller than the laser footprint (such as a leaf) or where the target has a vertical face (such as the edge of a building). Where a pulse is split there will be more than one reflectance return to the sensor. In some systems only the first pulse is recorded, however it is becoming increasingly common for systems to record up to five reflectance values (and some systems record the entire waveform). Such information is useful for, amongst other things, classifying laser points during post-processing. For example, areas highlighted as having multiple returns in urban areas will tend to be vegetation or building edges, and of the two vegetation will often produce more returns owing to the irregularity of the surface. More details regarding range equation dependence on a target area are given in Jelalian (1992).

2.5.3 Sources of Error in Data Capture

Airborne laser scanners are quoted as being capable of measuring heights to accuracies of 5-15cm and of planimetric positions to 0.5-1.0m accuracy (Behan et al. 2000; Maas 2000), depending on the type of terrain. However, despite increases in the accuracy of the system, research has indicated that during data capture the creation of error is inevitable. Sources of error in raw laser data have been attributed to inaccuracies in or problems with the GPS and INS systems used to obtain positioning information for the platform (Huising and Pereira 1998b; Lemmens 1997; Krabhill et al. 1995), or of the integration of the GPS and INS with each other or with the rest of the system. In addition, errors introduced during the matching of strips of data have been identified and quantified by Maas (2000) and Behan (2000). Errors from all sources have been categorised by Crombaghs et al. (2000) into 4 components: errors per point, errors per GPS-observation, errors per strip and errors per block. This research has highlighted the complexity of identifying the errors, which may come from a variety of sources. To further complicate matters the errors may either be omnipresent or can occur only under certain conditions (such as certain terrain types or weather conditions, Behan et al. 2000).

Errors in the GPS include those produced by interpolation, integer ambiguity, and iono- and tropospheric differences (vanEssen et al. 1997). INS related errors include gyroscopic drift and mis-alignment to the Earth's gravitational field. There are also errors produced by the laser itself, the magnitude of these are determined by the size of the laser footprint, the terrain type, and of course the calibration of the laser. The integration of these three systems is a further source of error, with the three principal sources being co-ordinate system transformations, transformations between sensors and synchronisation (Behan et al. 2000). Taking each of these in turn, GPS positioning accuracies can be between 3-10cm (Skaloud and Schwarz 1998) where high quality, differential carrier phase GPS receivers are used for baselines shorter than 30km. However, lower accuracies will be reported during missions where loss of GPS lock has occurred (most frequently during turns between strips). Errors

in the INS system relate to the estimation of rotation of the aircraft (roll, pitch and heading), and can relate to poor initialisation of the sensor or incorrect relative orientations between the INS and other sensors (GPS and laser). The integration of the GPS and INS works such that the GPS measurements correct the medium and long term INS instabilities, whilst the INS measurements increase the short term accuracy of the GPS measurements (Cramer 1999). However drift errors may occur, due to either incorrect modelling of the integration process or where low cost INS systems are employed which cannot attain the orientation accuracy required for high quality geo-referencing.

Error reduction strategies for data capture have been suggested by a number of authors, most of whom emphasise the importance of system calibration (Krabhill et al. 1995; Thiel and Wehr 1999). Other reduction strategies include careful mission planning (as advocated by Behan et al. 2000), such as ensuring that the aircraft is sufficiently close to the GPS base station. This is important, as increased distances from the base station cause degradation of the GPS data (Bean and Ferguson 2003). These authors also advocate the collection of static GPS data both at the start and at the end of the mission which can help in the resolution of ambiguities caused by the loss of lock and cycle slips. Krabhill et al. (1995) propose further strategies in mission planning which can reduce error, these include specifying the maximum banking angle for the aircraft between flight lines as this reduces the possibility of loss of lock (which can occur where the angle of elevation exceeds the angle of elevation of the satellites above the horizon). Low banking angles are also more suitable for INS efficiency as the effects of gravity (on which the gyroscopes depend) can be temporarily suspended if the platform tilts too much. Finally, the timing of the mission and the associated positioning of satellites can also affect the amount of error introduced during data capture.

Despite careful mission planning and calibration of instruments, some errors will still be present in the data (Behan et al. 2000). Overlapping data strips, together with external information and a mathematical model of the errors, are used to reduce

or eliminate these errors. Flight missions should be planned so that strips of data overlap. The size of the overlap region depends upon the flying parameters (such as scan angle and flying height) and can range from a few tens of metres to 1km (Behan et al. 2000). Many authors have proposed methods for matching data strips, some methods require the rasterisation of the irregular raw data, whilst others use the data in a TIN (e.g. Maas 2000) and propose a self-calibrating least squares adjustment to correct the measurements. Regardless of the method used, the identification of errors in laser strip matching is relatively straight-forward. Where there are no errors present then the strips will merge exactly with no ridges or gaps along the lines of overlap. They will also match exactly with ground truth data of equal or higher accuracy. However, as shown by Crombaghs et al. (2000) some adjustments will always have to be made. This requires that the magnitude of errors be quantified. Behan et al. (2000) and Crombaghs et al. (2000) differentiate between two types of errors: absolute offsets and relative offsets. The first can be corrected for by measuring the location of a point of known co-ordinates in all strips in which it is visible, the latter is calculated by measuring the discrepancies between tie points in the overlap between two or more strips.

2.6 Supplier Processing

Data suppliers are usually responsible for preliminary calibration and post-processing of raw data. These procedures are discussed briefly here as they are potential sources of error which should be included in any error budget calculation.

Prior to data capture, all of the system components must be mounted and calibrated. This includes testing for any systematic errors, such as time dependent drifts, or any atmospheric or ionospheric effects. The relations between the subsystems must be established, including the eccentricity vectors between the INS and the GPS, and the ALS and the INS. Time synchronisation parameters must also be established between the GPS and INS units, and between the GPS and the ALS. Following

data capture the coordinates for the laser point must be calculated. In ALS there are three coordinate frames used: aircraft body frame, INS reference frame and the Earth-fixed reference frame, these are shown in Figure 2.7. The aircraft reference frame (ARF) moves with the aircraft, with a rotation centre focused on the scanner mirror. In this system X_{ARF} represents the inflight direction, Y_{ARF} points to the right, and Z_{ARF} points down. This reference frame is linked to the scanner. The INS reference frame (IRF) is a local frame, the origin of which is a defined point within the sensor. The frame moves with the aircraft, always following the plumbline. The third reference frame used is the Earth Centred Earth Fixed Reference Frame (ECEF) which is fixed to the reference ellipsoid on Earth (WGS84), where the origin is defined as the centre of the Earth. In terms of axis orientation, X_{ECEF} is in the equator plane through the Greenwich Meridian, Y_{ECEF} is in the equator plane 90 deg east of Greenwich, and the Z_{ECEF} points northwards, parallel to the North Pole. In preliminary supplier post-processing, i.e. that immediately following data capture, some transformation is required between these reference frames in order to produce the final coordinates for the laser point. This is done using a 6-parameter transformation for each point from one frame to WGS84 using defined translation and rotation vectors. The general laser equation (see below) defines the position of a laser point in space, in the example below only (X) is defined for simplicity.

$$
\begin{aligned}
X_{LP}^{ECEF} = {} & X_G^{ECEF} + \Delta X_G^{ECEF} + R(0, \frac{\pi}{2} + \phi_G^{ECEF}, -\lambda_G^{ECEF}).R(\Delta\eta, \\
& \Delta\zeta, 0).R(r + \Delta r, p + \Delta p, h + \Delta h).R(m_r, m_p, m_h). \\
& (R(a + \Delta a).s_a, (b + \Delta b).s_b, 0). \\
& \begin{bmatrix} 0 \\ 0 \\ (l + \Delta l).s \end{bmatrix} - t_I^{ARF} - \Delta t_I^{ARF} - t_G^{ARF} - \Delta t_G^{ARF} \qquad (2.7)
\end{aligned}
$$

where:

X_{LP}^{ECEF} is the X coordinate of the laser point ΔX_G^{ECEF}

ΔX_G^{ECEF} is the GPS offset

R is the slant range from the aircraft to the ground

$\Delta\eta, \Delta\zeta, 0$ is the deflection of the geoidal normal

Δr , $\Delta p, \Delta h$ are the INS attitude offsets

m_r, m_p, m_h are the INS-ALS mounting misalignments

$\Delta a, s_a$ is the cross-flight scanner offset and scale

$\Delta b, s_b$ is the in-flight scanner offset and scale

$\Delta l, s_l$ is the laser offset and scale

Δt_l^{ARF} is the INS eccentricity offset

Δt_{ARFG} is the GPS atenna eccentricity offset

Source: Lindenberger et al. (2003). Further, more detailed, analyses of the errors produced during data capture can be found in Measures (1992).

The above equation outlines the formal procedure for transforming coordinates between the various coordinate systems, and combining these with the measured laser range to produce the full coordinates for each laser point, given in WGS84. The equation defines how the GPS coordinates (taking into account the GPS offset (X_G^{ECEF}) and the deflection of the geoidal normal at range R ($R(\Delta\eta, \Delta\zeta, 0)$) are multiplied by the values of the roll, pitch and heading as measured by the INS, and offsets, (in IRF coordinates) $(R(r + \Delta r, p + \Delta p, h + \Delta h).R(m_r, m_p, m_h)$. Mounting misalignments are also taken into account (R(m_r, m_p, m_h). The point is then rotated by the scanner angle (including any offsets $((R((a + \Delta a).s_a, (b + \Delta b).s_b, 0))$. Finally,

the position of the laser ($\begin{bmatrix} 0 \\ 0 \\ (l+\Delta l).s \end{bmatrix}$), the INS ($t_I^{ARF} - \Delta t_I^{ARF}$), and the GPS antenna ($t_G^{ARF} - \Delta t_G^{ARF}$) are all accounted for.

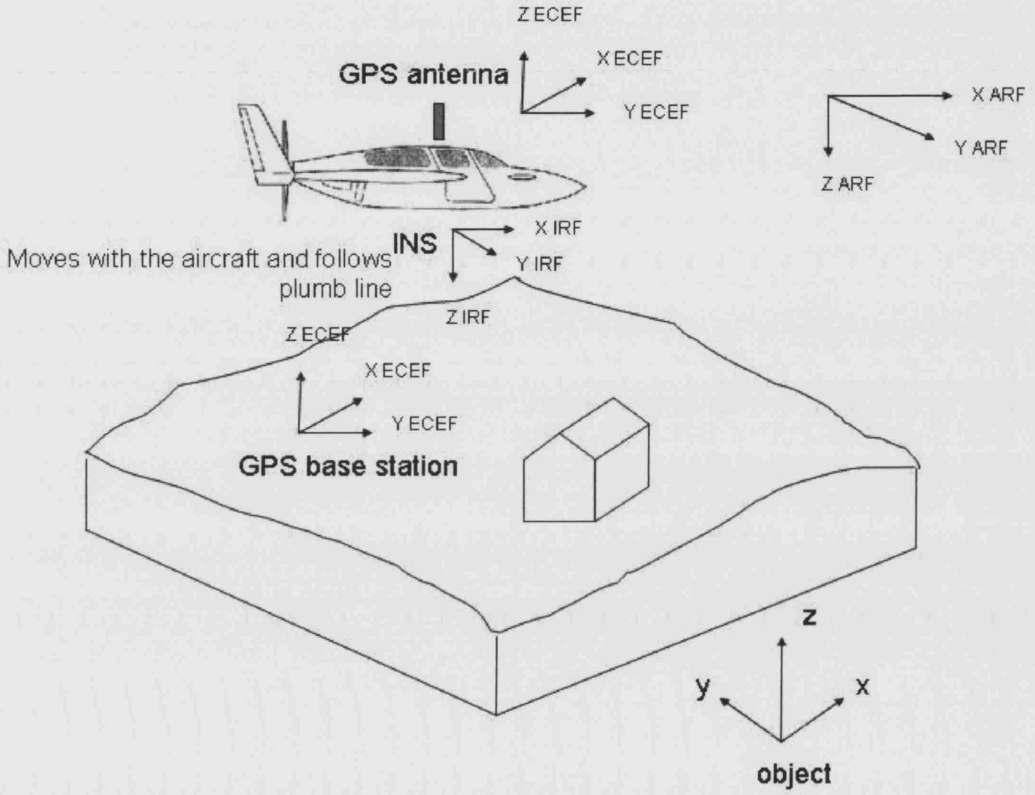

Figure 2.7: Airborne Laser Scanning Reference Systems (after Lindenberger et al 2003)

Such transformations may introduce error into the laser point during this preliminary post-processing, this and other potential sources of error during data capture are highlighted in the error budget equation which is introduced later in this chapter.

The second process conducted by data suppliers is that of strip matching. Laser points are captured in parallel areas or strips in accordance with the defined flight path. Following the appropriate transformations to determine the coordinates of the laser points within each strip, the whole strips need to be aligned in order that they match. The technique of strip matching determines the transformation

parameters which are required to minimise the height differences between adjacent strips. Chapter 3 details some of the methods for strip matching and alignment. Suffice to note here that all of the techniques involve some manipulation of the laser point coordinates, and so may introduce error within the post-processing flowline.

All of the sources of error discussed in this chapter are introduced during either pre-flight calibration, data capture or preliminary post-processing. As will be shown in Chapter 3, these errors have been previously studied and quantified. There are three different types of errors which may occur in an ALS dataset (Lindenberger et al. 2003): blunders, systematic errors and random errors. Blunders are simple mistakes, which are often caused by human carelessness. They are unpredictable, however they are reasonably easy to identify and eliminate from the dataset during post-processing. Systematic errors occur according to some definite pattern and are usually caused by miscalibration of the system, again these are relatively easy to identify. Random errors are caused by the inherent incapability of the instruments themselves to make exact measurements. They are also caused by uncontrollable variations in the operating conditions (Lindenberger et al. 2003). They tend to be very small in magnitude, and can be of any sign, furthermore, this type of error can never be overcome despite repeat measurements. Random errors appear to occur in an irregular nature, however when there is a large number of observations, random errors in measurements conform to the normal laws of errors, and so to some extent can be predicted.

Given what we currently know about errors in ALS data capture and preliminary post-processing it is proposed within this thesis that it would currently be possible for data suppliers to use this existing information to process all of these estimated errors within one equation (derived here for the first time) to define the error budget of each laser point, and by inference the dataset as a whole, such that:

$$\varepsilon_{LP} = f(\varepsilon_{cal}, \varepsilon_{GPS}, \varepsilon_{INS}, \varepsilon_{mm}, \varepsilon_{s}, \varepsilon_{R}, \varepsilon_{T}, \varepsilon_{sm}) \tag{2.8}$$

where:

ϵ_{LP} = the estimated error

ε_{cal} = error in pre-flight system calibration

ε_{GPS} = error in GPS measurements

ε_{INS} = error in INS measurements

ε_{mm} = error produced by mounting misalignments

ε_{s} = error in scanner angle

ε_{R} = error in Range measurement

ε_{T} = error in transformations from coordinate systems into common reference system

ε_{sm} = error introduced by strip matching

Where: ε_{GPS} is defined as

$$\varepsilon_{GPS} = f(\varepsilon_{GT}, \varepsilon_{G}^{ECEF}, \varepsilon_{A}) \tag{2.9}$$

where:

ε_{GT} = error in GPS clock

ε_{G}^{ECeF} = error due to GPS offsets

ε_{A} = error in GPS antenna eccentricity offset

and ε_{INS} is defined as

$$\varepsilon_{INS} = f(\varepsilon^{ARF}, \varepsilon_r, \varepsilon_p, \varepsilon_h) \tag{2.10}$$

where:

ε^{ARF} = error in INS eccentricity offset

ε_r = error in roll measurement

ε_p = error in pitch measurement

ε_h = error in heading measurement

$$\varepsilon_{mm} = f(\varepsilon_{vI}^{ARF}, \varepsilon_{vG}^{ARF}) \tag{2.11}$$

where:

ε_{vI}^{ARF} = error in calculation of vector from ALS to INS

ε_{vG}^{ARF} = error in calculation of vector from INS to GPS antenna

$$\varepsilon_T = f(\varepsilon_{tr}, \varepsilon_{rm}, \varepsilon_{IRF}, \varepsilon_{ae}, \varepsilon_{alp}) \tag{2.12}$$

where:

ε_{tr} = error in translation vector

ε_{rm} = error in rotation matrix

ε_{IRF} = error in transformation from origin of IRF to laser point (LP) into IRF

ε_{ae} = error in transformation of the antenna eccentricity (ARF) into IRF

ε_{alp} = error in transformation from IRF to ECEF of the vector from the GPS antenna to the laser point

The propagation of the errors through this processing chain is likely to be extremely complex, and most probably will be multiplicative.

As outlined in Chapter 1, for the majority of applications, data-users will require additional post-processing of the raw points in order to produce height models such as those shown in Figure 2.8.

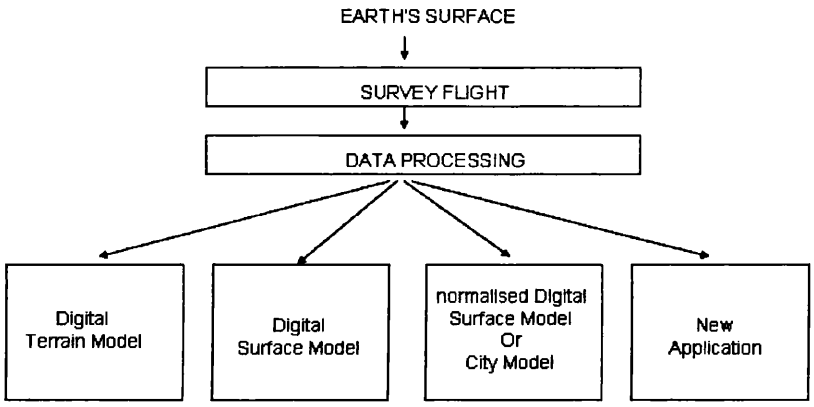

Figure 2.8: Modelling from Laser Data

This end-user post-processing also introduces error into the laser data and as such forms an important component of any full error budget calculation. However, as identified in Chapter 1, the procedures within this end-user post-processing workflow have not been formally quantified in any previous research. The research presented in this thesis focuses on the analysis of the errors introduced during end-user post-processing, the parameters for quantification are presented in the extended error budget equation below:

$$\varepsilon_{LP} = f(\varepsilon_{cal}, \varepsilon_{GPS}, \varepsilon_{INS}, \varepsilon_{mm}, \varepsilon_{s}, \varepsilon_{R}, \varepsilon_{T}, \varepsilon_{sm}, \varepsilon_{ds}, \varepsilon_{f}, \varepsilon_{s}, \varepsilon_{rcon}) \qquad (2.13)$$

where:

ε_{ds} = error introduced by data structure

ε_{f} = error introduced by filtering

ε_{s} = error introduced by segmentation procedure

ε_{rcon} = errors introduced by reconstruction of object vectors.

This thesis presents a quantification and analysis of the following parameters: ε_{ds}, ε_{f}, ε_{s}. The form of the function is not defined, given that this would require detailed statistical analysis of the propagation of error from supplier processing through to user post-processing which was beyond the remit of the investigation, but is recommended for further work in Chapter 11.

2.7 Chapter Summary

This chapter has presented the fundamental principles of airborne laser scanning, including descriptions of the system sub-components and how these are used to determine range information. Potential sources of error are presented within one equation. The inclusion of post-processing errors within this equation has not previously been proposed. This equation underlies many of the investigations presented in this thesis, each of which attempts to quantify one of the post-processing parameters.

The following chapter presents the context from which the thesis has arisen.

Chapter 3

Research Context

3.1 Chapter Introduction

This thesis is concerned with the modelling of elevation and error from ALS data, and draws primarily from three areas of previous research: urban modelling (including representations of population and climate surfaces), terrain modelling (which has traditionally focused on non-urban areas), and remote sensing. The thesis, in part, attempts to integrate techniques developed in each of these three fields (Figure 3.1) in order to examine the creation methods for urban elevation surface models from ALS data.

The chapter is structured in accordance with the three areas identified above, and begins with a review of methods for representing urban surfaces, including a discussion of appropriate methods for data modelling. Secondly, previous work detailing the modelling of elevation is reviewed. Many of the techniques and theories underlying this large body of research are relevant to urban elevation modelling, and some are directly relevant to this thesis. Finally, a review of urban remote sensing is offered, including an overview of the applications and problems currently faced in this field. The post-processing and modelling from laser data forms the major part

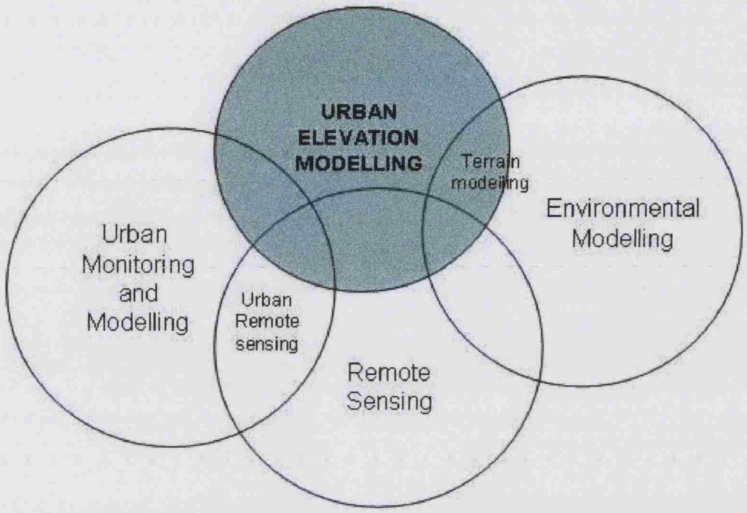

Figure 3.1: Diagram depicting research context for thesis, and outline of issues discussed in this chapter

of this section, and it is here that many of the specific research problems, which are investigated in this thesis, are introduced. The chapter concludes with a discussion of the modelling of error in elevation models, and the sources of error in ALS data.

In summary, this chapter discusses why urban elevation modelling is important, why ALS data should be used in the creation of such models, and how the accuracy of these models may be measured.

3.2 Representations of Surfaces

One of the most important choices for the ALS modeller is that of representational form for the data during processing. The individual laser points in the dataset represent discrete elevation values, and it is the job of the user to determine whether these should be processed as discrete entities or whether the intervening gaps should be interpolated across in order to create a continuous surface model. The choice depends not only on theoretical issues (such as which is the most suitable model for

representing urban form) but also on the availability of algorithms and software for processing data - some of which may require a specific data structure.

Chorley and Haggett (1965) noted that representational mode is an outcome of choice, convention, and convenience. In ALS modelling, it is often convenience which is the deciding factor when choosing either a field or object model. This is largely driven by the requirement of various software packages for input data to be modelled in a certain way. To examine the theoretical consequences of using different representations a wide range of literature was examined for this research, including an examination of representations of urban surfaces for topics such as climatology (Fan and Sailor 2005; Sailor and Lu 2004), air pollution (Scaperdas and Colvile 1999; Patil et al. 2003) and population studies (Dodge et al. 1997; Martin 2002). A number of parallels may be drawn between such investigations and representations of urban heights using ALS data as many of the problems facing surface modellers (regardless of discipline) are similar. Examples include theoretical issues regarding the nature of the objects being modelled, as when representing distributions of unique human individuals as continuous population surfaces. In such models the human beings which are modelled are discrete objects and yet are frequently displayed as continuous fields. It is highly likely that errors will be generated by the spatial object transformations that are invoked (Martin 1991, 1996). In a similar way, in ALS modelling the discrete points only represent the heights at the particular (x, y) location and a small ellipse surrounding this as defined by the laser footprint. However, the points are frequently used to represent the height over a much larger area, and often are represented in terms of a continuous model. Points representing discrete entities (buildings, street lights, bus shelters) clearly cannot and should not be used to infer the height of the surrounding ground surface.

There are also parallels in terms of the nature of sample points from which continuous representations are interpolated. Urban climate and air pollution models are often created from samples taken at pre-defined stations, despite the fact that the

phenomena being modelled are moving in time and space, often at long distances from the data collection point. Clearly, this may cause some large differences in the prediction of urban climate/air pollution and the observed levels. In a similar way the values in ALS datasets are also discrete measurements of a varying phenomenon, in this case elevation, and some of the same problems apply.

There are limits to all types of urban models in terms of scale and resolution, and whilst there may be different reasons for this (such as preservation of anonymity in population modelling) the effect is similar in terms of the introduction of errors into the models. For ALS modelling, the location of the sample points is defined by the flying height, aircraft speed, scanner oscillation, and by the scanner mechanics. Using different settings/equipment would produce very different samples, and consequently may produce different surface representations (Raber 2003). The point density with which ALS data may be collected is increasing rapidly, with advances in scanner mechanics, but the fact that some interpolation is inevitable in order to model intervening locations means that there is still some element of error, and uncertainty, within the model, as is the case with any representation (Longley et al. 2001, 2005).

In summary, there are a number of parallels which can be drawn between the modelling of urban elevation surfaces from ALS data and the modelling of other urban phenomena. Whilst the focus of this thesis is on understanding the nature of errors in ALS elevation models, the intrinsic links with other urban surfaces (such as those outlined above) suggest that the methodology used in this investigation may be, at least in part, of generic importance in interpolating between points in a range of other applications.

Whilst this discussion has focussed largely on theoretical issues, in many respects the choice of which representational model to use for ALS data actually has less to do with ideals and more to do with convention. This is shown below with reference to previous studies of elevation modelling.

3.2.1 Elevation Modelling

This thesis is primarily concerned with the issue of how ALS data are processed, stored and represented. These issues are inextricably linked, particularly in the case of ALS data, as the choice of storage model (continuous or discrete) largely dictates how the data will be processed.

The acquisition, storage, and representation of topographic information has been an area of active research for more than twenty years (see, for example, Mark 1978; McCullagh 1988; Fisher 1991; Wood 1996; Frederiksen et al. 1998). In terms of data storage, elevation may be sampled at discrete points or may be represented as averages (which tend to then be modelled as though they were discrete points). In practice objects (in this case the discrete points) are usually vectorised whilst fields are rasterised. However there is no necessary correspondence.

The theoretical debate regarding the choice of data model (field or object) to use in the creation of any elevation model is a fundamental one, and depends largely on the application for which the model is intended. Traditionally, terrain height has been conceived of as a continually varying surface (Burrough 1986), which may be stored as either a grid, as a TIN, or as a contour map. Such digital elevation models are the primary data used in analysis for a number of applications including hydrology, biology (Moore et al. 1991), and geomorphology (Gardner et al. 1990). For the purpose of numerical analysis and modelling, surface elevation is usually represented using either a grid (generally referred to as a gridded DEM) (e.g. Skidmore 1990) or a triangular irregular network (TIN) model. In the gridded model (DEM) variation is described by determining the variable's value within each rectangular cell and spatial variation within each cell is ignored. Gridded DEMs are usually stored as gridded matrices of elevations, and the associated grid resolution (Kumler 1994). In the TIN model, only critical points which define local discontinuities are recorded and as such these points will tend to be irregularly spaced. These points are connected together in a network of edges that form space-filling, non-overlapping

triangles (Kumler 1994). The points are most usually connected using Delaunay triangulation which triangulates using the centres of neighbouring Thiessen polygons (Thiessen 1911; Delaunay 1934). Other, more complex procedures are available, some of which incorporate additional information such as whether the edge is a contour or a breakline.

It is common to represent elevation, including ALS data, as a continuous model. Indeed, a number of products may be derived from either the grid or a TIN model, these include contour maps, viewshed and line of sight maps, maps of derivatives such as slope, aspect, convexity and concavity, and shaded relief maps. There are merits of both the DEM and the TIN structures as models of continuous form, and these have been intensively investigated in the past. For example, Kumler (1994) aimed, first, to compare the relative efficiencies of TINs and gridded DEMs as representing a wide variety of terrains at uniform scale, and, second, to ascertain how best to select the vertices for a TIN. Kumler (1994) demonstrated the advantages of the TIN where terrain is rapidly changing. In such situations the TIN can include many points, as its resolution is variable, whilst also including only a few points where the surface is relatively uniform. However, Kumler (1994) also showed how these advantages are countered by complexities in both storage and manipulation, and concluded that contour based TINs, in his study, were not more efficient than gridded DEMs.

Despite the popularity of DEMs, there are a number of conceptual problems with DEMs as models of surface form, many of which were highlighted by Wood (1996). Firstly, despite being a model of continuous surface elevation, the gridded DEM is actually a set of regularly spaced discrete elevation values, and as such Wood (1996) states that the accuracy of the DEM is essentially a function of true terrain roughness, DEM resolution (which was also examined by Balce 1987), and of the interpolation method used to create the DEM (Wood and Fisher 1993; Lam 1983). The implications of using different interpolants for deriving intervening elevation values is examined in Chapters 6 and 7 of this thesis. The second conceptual prob-

lem, which is intrinsically related to the first, is that there is a lack of consistency in the definition of what each elevation value represents within the DEM. In some models, the elevation value relates to the centroid of the grid cell, whilst in others it relates to the elevation value at the grid node. For any overlay analysis it is imperative that this be ascertained.

Continuous models of the urban surface are frequently used for applications analysing surface derivatives such as gradient and aspect. These models are used in planning for viewshed analysis and line-of-sight investigations. They are also used in flood predictions, and run-off calculations, both of which require a continuous surface form. However, many aspects of the urban environment are more intuitively conceived of as discrete entities, which occupy an otherwise empty geographical space. Zhang and Goodchild (2002) cite the applications area of urban planning and management, in which the representation of complex urban landscapes requires the use of well-defined objects to represent entities such as buildings, the road network, and utilities networks. An active area of research in urban elevation modelling has been the identification of which elements of the urban physical environment should be included in discrete models. Zlatanova (2000), Gruen and Dan (1997), Kofler and Gruber (1997) and Tempfli (1998) all identified buildings as being the most important element of the urban environment, and it is therefore no surprise that most object based models of urban areas have focused on buildings. However, Fuchs (1996) suggested that additional objects of interest, following buildings, should also be incorporated into models, and should be grouped as follows: vegetation, traffic network, public utilities, and telecommunications. In fact Fuchs (1996) drew no distinction between the importance of buildings, vegetation or traffic network, conclusions which have been corroborated by Ranzinger and Gleixner (1995) and Dahany (1997).

In summary, elevation data have traditionally been structured and modelled as continuous fields. However, for many applications urban elevation is more conveniently conceived of in terms of an assemblage of discrete entities or objects. General con-

sensus has developed over the years about which data structure should be used for particular applications, however in the case of urban surface modelling the choice is less clear. Zhang and Goodchild (2002) state that when the phenomena are considered to be real or static entities then the discrete object model should be used, and where phenomena are ubiquitous, a continuous model should be adopted. In the case of elevation, and particularly urban elevation, both approaches may be suitable. For example, whilst the elevation of the urban surfaces may be largely continuous, this surface is often disrupted by discontinuities such as building edge, and roof ridges, and frequently contains discrete entities such as roads, streets, buildings, and parks. Whilst, it may be more theoretically sound to treat urban elevation as a collection of discrete objects, in many ALS processing methodologies it is often not possible to infer the morphology of urban features (such as buildings) without first creating a surface model which may actually highlight surface discontinuities.

3.3 Urban Remote Sensing

The second body of literature from which this thesis draws is that of urban remote sensing. This section highlights some of the problems facing the remote sensing of urban areas, and justifies the choice of ALS data as the focus for this research.

Urban areas have long been the focus of research (Welch 1982; Forster 1983; Aplin et al. 1997; Barnsley and Barr 2000; Barr and Barnsley 2004). This is unsurprising given that the majority of the world's population inhabit such areas (Mesev 2003). Remote sensing can offer valuable sources of information for the monitoring of the growth of urban areas (such as Grey and Luckman 1999). For example, remotely sensed data may be used in predictive models to estimate levels of sustainable expansion for these areas, this is measured in terms of the levels which may be tolerated by both society and the environment, including studies such as the exploitation of natural resources, associated air pollution, and climatic changes (Yeh and Li 2001). Other uses of remotely sensed data include planning issues such as housing demand,

waste disposal and highways planning (Ogrosky 1975; Michalak 1993; Couloigner and Ranchin 2000). In terms of urban planning and highway engineering, remotely sensed data have principally been used to provide context and to convey additional information about the human landscape during analysis (Carlson 2003). The use of data in an essentially contextual role is unsurprising given the complexities of the urban landscape. Lower resolution datasets (such as those from satellite sensors including IKONOS, SPOT, METEO-SAT, MODIS, DSMP, and Landsat) have been used for context in a variety of urban planning studies. Examples include analysis of urban development (Batty and Howes 2001), land use pattern (Barnsley and Barr 1996; Barnsley et al. 2003) and change, including predictive modelling of this (Méaille and Wald 1990; Pathan et al. 1993), urban sprawl, modelling of urban surface run-off (Gillies et al. 2003), the modelling of urban climatology (Hammer et al. 2003) and the urban heat island (Nichol 1988; Eliasson 1992; Voogt and Oke 2003; Gomarasca et al. 1993), population modelling (Longley and Mesev 2003) and the more general modelling of the quality of life (Donnay and Nadasdi 1992; Weber and Hirsch 1991; Richman 1999). There has also been a considerable amount of research conducted which takes these physical models and analyses the spatial configuration of structures in order to understand the human behaviour which brought about the spatial patterns (e.g. Geoghegan et al. 1998).

As availability and capture accuracies of remotely sensed data increase, they are more commonly being used to create models of the physical urban environment. Donnay et al. (2003) noted the importance of this recent focus, stating that 'the extraction of planimetric and 2.5/3D features from remotely sensed images.....constitutes a very dynamic research domain in urban remote sensing'. IKONOS and QuickBird imagery has been used with some success for feature extraction (e.g. Sohn and Dowman 2002a). The former is now commercially available at a spatial resolution of 1m and the latter at 0.6m. These data have been used in a variety of studies and recent work by Sohn and Dowman (2002a) has reported promising results. However in their study only large detached buildings were of interest, and as such the applicability of this technique for the entire city environment is therefore limited

perhaps emphasising that inferring the detailed morphology of individual buildings remains a difficult task and one which tends to be error prone. The benefits of using Quickbird imagery were also examined by Holland and Marshall (2003). The authors found that even in well-mapped countries there was enormous potential for Quickbird as a data capture source.

Despite the potential of satellite remote sensing as a tool for understanding human behaviour and for creating physical models of the urban surface, there are a number of limitations associated with this capture technique which are discussed briefly here. Urban surfaces remain one of the most challenging environments for inferring the detailed morphology of land cover and land use from imagery (Mesev 2003). This is mainly because of their complexity and mix of both artificial and natural structures. For satellite imagery, one of the greatest problems, despite an increase in spatial resolutions available, is that artificial structures are not easily modelled from pixels (which may represent $400m^2$ in the case of SPOT-HRV XS, or $900m^2$ for Landsat TM). Thus for lower spatial resolutions pixels may represent a mix of surfaces rendering classification difficult, whilst for higher resolutions pixels may be smaller than urban objects and as such may introduce noise into a classification (Atkinson and Curran 1997). Carlson 2003 states that the resolution of satellite data has been too coarse for discriminating built environment features such as houses and roads from natural surfaces in the urban landscape.

There should be a distinction made here between any technical advance in measurement instruments and the problems of inferring morphology at almost any scale of measurement. This suggests that the solution to the problem cannot be purely technical and restricted to measurement instrument development. Presumably the optical capacity of measurement instruments has a finite limit and thence it will be necessary to develop better inferential procedures.

The extraction of urban features using photogrammetric techniques has also been attempted. Photogrammetry is the science and technology of obtaining spatial mea-

surements from photographs, and can be used with both aerial and satellite images. Typically, procedures involve the measuring of distances, areas and elevations. Studies by Baltsavias and Gruen (2003) have demonstrated the lower performance of satellite imagery (even high resolution) for feature extraction compared to digital aerial photography, highlighting why urban feature extraction has been dominated by the use of aerial photography.

Numerous techniques for object recognition and feature extraction from aerial photographs exist (many of which are presented in Nixon and Aguado (2002)). Images may be processed to facilitate the extraction of edges or homogeneous regions. These edges are subsequently combined using geometric and/or perceptual rules in order to complete the object description. Bellman and Shortis (2000) have described how in some cases edges are matched to models of generic objects. This, in fact, implies a cognitive approach to vision, where the recognition task is largely performed as a cognitive rather than a visual process. Results from aerial photogrammetry are often the most detailed and offer an attractive methodology for urban model creation. There is a clear difference in the degree of operator intervention that is required for modelling from different data sources, and photogrammetry requires a very high level of intervention. This makes it a time consuming, expensive, and reasonably inefficient way of producing models which is simply not practicable over geographically extensive urban areas. Having said this, more recent papers, such as Gerke et al. (2001), have used an approach for the detection of buildings which reduces the amount of intervention required. In this study knowledge about the surroundings of each building were employed in order to support the detection of individual buildings.

However, aerial images are also costly to collect and to interpret, and consequently there has been a recent trend towards using less expensive, less labour intensive sources from airborne platforms such as Synthetic Aperture Radar (SAR) and Laser Scanning (ALS). Improved planimetric and vertical accuracies in these airborne sources have facilitated their popularity, as have the increasing number of companies

supplying such data and supplying them at competitive prices (Flood and Gutelius 1997; Flood 2001).

SAR offers a number of advantages over aerial photography, firstly that it is an active technique. Active imaging technology is that in which the sensor provides its own source of energy, which means in practice that SAR is independent of sunlight, and can therefore operate in both day and night. All radar techniques are cloud penetrating, which is a significant advantage over optical imaging. This is because radars use the microwave region of the electromagnetic spectrum and therefore have long wavelengths, allowing them to also penetrate many atmospheric particles. SAR is a technique which synthesises a very large aperture from a series of radar pulses transmitted by a small antenna. This is achieved using coherent radiation and by accounting for the Doppler phase shifts of the returned pulses. An extension of the basic SAR principle is interferometry (IfSAR or InSAR) which can be used to determine the third dimension of an object's position. In IfSAR, two images are used, with very small differences in their collection geometries. Recent methodologies for extracting elevations from such images are described in Hepner et al. (1998), Gamba et al. (2000), and Toutin and Gray (2000). SAR images tend to have a relatively coarse resolution which makes the differentiation of complex objects in close proximity difficult - particularly in the urban environment. There are also problems for side-looking radars in areas of steep slopes and complex urban areas due to foreshortening (Grey and Luckman 1999). Balz and Haala (2003) have also noted problems such as speckling and extreme reflections which make the use of INSAR problematic for modelling urban areas.

In light of the limitations of SAR, Baltsavias and Gruen (2003) proposed that the future of urban physical modelling may lie in Airborne Laser Scanning (ALS), for detailed and accurate identification and mapping of terrain, buildings, and other features, in both the vertical and horizontal planes. As shown in Chapters 1 and 2 modelling from ALS data involves a number of specific processes. Previous research relating to these processes are reviewed in the following section.

3.4 Previous Modelling Techniques Using Airborne Laser Scanning Data

3.4.1 Filtering Strategies

Filtering is the abstraction of the bare earth from a laser scanner dataset, and involves the removal of any point or cell from the dataset which does not represent terrain elevation. Different filtering methods require that data are formatted either as discrete points (although these are often held together as a continuous TIN model for processing), or that the raw points are reformatted onto a regular grid and processed as a continuous surface. The different approaches introduce different amounts of error into the derived models and this requires quantification if users are to make informed decisions about the optimal choice of processing methodology for any given application. There is a need for further research in this area.

There are numerous filtering techniques for each of the data format approaches. Most are based on the premise that the characteristics of terrain (such as slope, roughness, form, curvature, etc.) are bounded, and that surfaces which fall outside these bounds are not terrain (Sithole 2003). Many methodologies use discriminant functions to determine whether a candidate point represents elevation of the terrain or an object above this. Some of these algorithms work on a comparison of one point with another, others on a comparison of one point to numerous points, and still others on a multiple points to points basis. Many techniques also seek to measure discontinuity in the surface, with some computing height differences, others using slope. Those methods which specifically use the raw points may use measures such as the shortest distance to a TIN facet. Using this discontinuity information, a candidate terrain surface can be identified and used, with a defined threshold, to assess whether neighbouring points or cells should also be part of this terrain surface. The threshold is sometimes a buffer zone around either a horizontal plane (such as Wack and Wimmer 2002 who used this on gridded ALS data in their filtering method) or

a buffer around a defined slope representing the candidate terrain surface. In slope based filtering the slope, or height difference, between two points is measured. If the slope is above a certain threshold then it is assumed that the higher point is part of an object above the terrain surface. In both block minimum and slope based filtering the buffer zone around the plane defines a region in 3D space where terrain points are expected. Problems in identifying candidate bare earth points may occur using these methods for modelling in steep terrain areas.

Examples of slope based filters include those of Elmqvist (2001), Pfeifer et al. (2001a), and Brovelli et al. (2002). In the Elmqvist (2001) method a polynomial is fitted to the lowest points in the gridded dataset, and this deformable surface is then used to estimate the surface. The method of Pfeifer et al. (2001a), termed robust interpolation, computes an approximate surface based on the lowest points in the dataset once gross errors and outliers have been removed. The distance between each point and the computed surface is measured, and assigned a weighting in accordance with this distance. The surface is then recomputed, taking into account the weightings of the points. A point with a high weight attracts the surface whereas a point assigned a low weight will have little influence on the form of the derived surfaces. The Pfeifer et al. (2001a) and Brovelli et al. (2002) methods are based on the original method written by Kraus and Pfeifer (1998), which was designed to filter vegetation. Adaptations have rendered it suitable for filtering buildings in the urban environment. The method has been further developed by Schinckler and Thorpe (2001) to identify breaklines in point clouds and to categorise the surfaces identified. Sithole and Vosselman (2003) found that the Pfeifer et al. (2001a) filtering algorithm produced good results particularly in the removal of very large and very small objects: however, slight problems were noted in the removal of very low objects, and those with high complexity. As one of the most successful algorithms in the Sithole and Vosselman (2003) comparison, the Pfeifer et al. (2001a) algorithm was chosen as one of those to be investigated within this study. The adaptations made to this algorithm for use in this thesis are detailed in Chapter 8.

Axelsson (1999) developed an alternative to the slope based filtering approach in which he derived a terrain model from the raw points using the progressive densification of a TIN created from the points. His method works by defining an approximate terrain surface based on neighbourhood minima, and then computing the distance and direction to each of the points above the surface. Points within a defined distance range are accepted and are then used to recompute the TIN, and new thresholds are computed. The method iterates through this process until no more points are beneath the threshold. It has been noted previously that this algorithm is particularly efficient at dealing with discontinuities in the surface, and as such is thought to be a suitable method for surface modelling in urban areas (Sithole and Vosselman 2003). This method was compared to others by Sithole and Vosselman (2003), where it was shown to perform well although it was found that it was too heavily influenced by low points. On the basis of this favourable report the Axelsson (2001) algorithm was chosen as the second method to be studied within this investigation (see Chapter 8).

The idea of using a progressive densification of the TIN was also employed by Sohn and Dowman (2002b). In this method, the authors employed the dual strategy of both upward and downward densification of the TIN. Results from this algorithm appear successful (Sithole and Vosselman 2003), although there are noticeable problems in the elimination of low rise objects, and the general preservation of surface form. Other problems include the determination of the first level of the TIN, where corner points are chosen to represent the terrain, with the obvious problem that these points may actually be representative of object heights rather than terrain.

The use of morphological filters for both gridded data and raw data have also proved popular. These methods classify the points, or cells, based on the height values within a given search window. One of the most frequently cited of the morphological filters is that of Vosselman (2000). In this method a predefined structuring element (in the shape of an inverted bowl) is pushed vertically through the data from beneath the point cloud. The structuring element is centred horizontally on a laser point.

It is then raised until it encounters a point in the point cloud. If this point is not the one on which the structuring element was first centred, then the centering point is treated as not belonging to the terrain surface. In other words, a point is only accepted as terrain if it is the first point that the structuring element encounters as it is raised through the point cloud. Problems have been identified with these methods (Sithole and Vosselman 2003) where they are applied to areas of steep terrain, or in urban areas where points are aligned vertically along a building wall or along a steeply inclined roof structure. Vosselman and Maas (2001) have attempted to improve the performance of the algorithm in these regions, by using a slope map to identify the angle at which the structuring element should be moved through the points. Other morphological filters include that of Roggero (2001)(2002) who applied the technique to gridded data. In this method, the local slope is determined using a local operator to determine the lowest point in an area of interest. This low point is assumed to be part of the terrain.

There are many other filtering techniques which have been developed which are too numerous to review in detail here. Some of the more significant methods include the use of splines (Brovelli et al. 2002), the use of multiple laser returns and specifically the difference between the first and the last return (Altharthy and Bethel 2003; Steinle and Vögtle 2000), and the use of geometric characteristics of the data (Priestnall and Glover 1998; Priestnall et al. 2000). Recently, emphasis has been on the development of many of these algorithms so that they are suitable for either sloping terrain (e.g. Vosselman and Maas 2001) or the removal of complex structures on flat terrain (e.g. Masaharu and Ohtsubo 2002).

What is clear from this review is that there is a variety of filtering strategies for ALS data, and these either use a discrete modelling approach (e.g. Sithole and Vosselman 2003), a continuous modelling approach (such as the Wack and Wimmer (2002) technique which used gridded data), or a combination of both discrete and continuous modelling approaches (e.g. the Axelsson (1999) method which inputs and outputs discrete points, but processes these using a continuous TIN).

Comparisons of the success of various filtering algorithms have been limited to a handful of studies which include Huising and Pereira (1998a), Haugerud and Harding (2001), Tao and Hu (2001), and Sithole and Vosselman (2003). These investigations have tended to use global descriptors of error. Despite this, most previous studies have made a distinction between Type I and Type II errors. Type I error is the misclassification of terrain points, or cells, as objects. Type II error is the misclassification of objects as terrain. Of the aformentioned investigations, only Sithole and Vosselman (2003) offered an experimental comparison in which they compared the results of eight automatic algorithms for extracting the terrain heights in order to determine which was the best methodology for certain terrain types. Whilst offering one of the most comprehensive comparisons of filters to date, this study did not compare the spatial location of errors within the derived models, nor did it show how this error information could provide additional information for modelling or how this could be used as a decision support tool. In a separate study Sithole (2003) did acknowledge the spatial variance of the success of filtering strategies and proposed a method in which the regions of a laser point cloud were identified as being potentially problematic for a filtering algorithm. In this method the point cloud is sorted into surfaces which are then compared with each other to determine height difference between them. Where the height difference value is great, the surface is assigned a high score between the values of 0 and 1, where 0 is terrain, and 1 is an object. In this way the point cloud is classified not only as terrain or object, but the certainty with which these classifications are made (as implied by the score given). Sithole (2003) suggested that such measures can be used as part of the filtering strategy in order to drive the filtering algorithm, and identify regions where additional information (such as an external data source) may be required. In the research presented here a similar suggestion is made, although emphasis is on the identification of regions of errors rather than simply local differences in height.

3.4.2 Segmentation Techniques

In Chapter 1 segmentation was defined as the division of an image or dataset by the classification of neighbouring laser points, or pixels, having common characteristics which define them as being 'above ground'. The majority of segmentation techniques have focused on detecting building candidates in a region. This is not an easy task, as there is often ambiguity given that there are numerous other objects in the urban scene which have similar characteristics to buildings. For this reason a number of authors have opted to identify building candidates using additional sources of information to the ALS data. Brunn and Weidner (1997), Vosselman and Dijkman (2001) and Hofmann et al. (2002) all used ground plans to reduce the processing time and to circumvent the building detection problem. The use of this approach is based on the assumption that the ground plans are always available, and are accurate, both in the positioning of the buildings and the currency of the data. Discrepancies between the two datasets will create errors in the derived building model. Other datasets which have been used in conjunction with ALS data include multispectral reflectance images (as used by Haala and Brenner 1999). The vast majority of research into segmentation techniques has not used these additional datasets, largely because in many instances ancillary data are not available at the required accuracies, and as such the suitability of ALS data as the only source of information needs to be tested.

There are many methods for segmenting laser data, the most cited are mentioned below. The most common method of segmentation for raw laser points is by using mathematical morphology methods (such as Maas and Vosselman 1999; Vosselman and Dijkman 2001). In these methods planar faces are identified based on clusters of common aspect values in a TIN. Filin (2002) took a departure from this method of detecting planes within the point cloud and instead looked at extracting other homogeneous structures, with some success. Other authors have used related similarity measures to group points, such as height and aspect (e.g. Haala et al. 1998 and Morgan and Tempfli 2000). Some authors have also used artificial neural

networks in order to discriminate between buildings and trees in the urban environment (e.g Priestnall et al. 2000). Maas (1999) adopted an interesting approach for segmentation, based on texture measures (the local variations in height), with a claimed classification accuracy of some 98%.

It is difficult to comment on the relative success of the various segmentation techniques, as to date no published comparison exists. But what is clear is that filtering and the segmentation are inextricably linked, and the accuracy of *both* processes is extremely important, particularly where the results from these processes are to be used for further analysis or modelling. The output from segmentation is often used in the creation of 3D building models (e.g. Maas 1999; Maas and Vosselman 1999; Wang and Schenk 2000; Morgan and Habib 2002; Rottensteiner and Jansa 2002; and Altharthy and Bethel 2003). Given this, it is clear that the accuracy of the final model depends not only on the success of the reconstruction technique, but also on the accuracy of the segmentation procedure.

Whilst the modelling of error, and error propagation in ALS post-processing have been largely neglected in previous studies, techniques for measuring the errors in other types of elevation model have been the focus of research in a number of disciplines for some years. This work is relevant to the development of a methodology for modelling error in ALS models. The principal techniques for error modelling in elevation surfaces are reviewed below.

3.5 Measuring Error in Elevation Modelling

Modelling the shape of the Earth is not a simple problem, particularly in the case of urban elevations, and the irregular nature of elevation surfaces excludes the possibility of ever knowing the true elevation at any point (Kumler 1994). There is, nevertheless, a need to assess how well, or badly, elevation has been modelled, and this has given rise to numerous research projects in which the nature of error in

elevation modelling has been investigated. The issue of assessing the accuracy of DEMs has been widely researched in the past by photogrammetrists and mapping scientists (Hannah 1981; Li 1988; McCullagh 1988; Carter 1989; Li 1993; Li 1994; Wood and Fisher 1993). Despite this work, ascertaining the accuracy of elevation models remains problematic, because of the complex interactions between different types of error, spatial irregularities of elevation surfaces, and error complexity (Zhang and Goodchild 2002). In 1991 Fisher claimed that the spatial distribution of errors across any DEM and the factors affecting the distribution of this error are largely unresearched, and to some extent this lack of understanding still exists today. Indeed, Zhang and Goodchild (2002) recently stated that discussion of DEM-related uncertainty is still an endeavour of great significance.

Previous work has shown that there are many potential sources of error within elevation models. Errors may be caused by systematic or blunder errors in data capture, introduced during data processing (e.g. digitising, photogrammetry, or in the case of ALS data the reformatting of raw data: Fisher 1990). The accuracy of the elevation model is, therefore, partly a function of the conceptual limitations of the model and partly of the quality of the data provided (Wood 1996). Whilst the measurement of error in elevation models has been the focus of much attention, there has been a lack of consistency in the methodologies for assessing elevation model accuracy . Some authors have experimentally assessed accuracy by comparing values within the DEM against a corresponding set of control points which are used as a reference dataset (e.g. Lloyd and Atkinson 2002). Many authors have attempted to model the complexity of error within elevation models, although a major handicap has been the need for the collection of ground truth data (Monckton 1994). This handicap has directly led to several attempts at developing theoretical models of error, some of which are based on mathematical models of terrain (e.g. Polidori et al. 1991; Li 1993) and as such it is unclear whether the conclusions are applicable to real terrain.

Published DEMs are usually supplied with accuracy metadata which may be used

as a reference for users. Ordnance Survey, the National Mapping Agency of Great Britain, supply these as a single value of root mean square error (RMSE), whilst the U.S Geological Survey use both RMSE and LE90 . LE90 is an abbreviation which stands for Linear Error at 90%. The designation LE90 is used to quantify the elevation error in a DEM versus the real world or truth DEM. LE90 is the error range which would include 90% of the pixels within the DEM. Thus an LE90 of 40 would indicate that 90% of the pixels in a given DEM vary from the truth by 40m or more. However, whilst this may be a useful indicator of general accuracy, difficulties arise when users wish to assess the reliability of results based on the DEM, given the occurrence of error within the DEM. For such analyses alternative measures of modelling and communicating error information are required. This is explored further below.

Methods designed to measure error can broadly be divided into those used for discrete models, and those applied to continuous surfaces.

3.5.1 Error in Discrete Models

In the discrete object approach there are five types of quality which refer directly to error, according to Zhang and Goodchild (2002): positional error, attribute error, logical inconsistency, incompleteness, and lineage. Of these, the most relevant to this study is positional error, particularly in the z dimension. This may be assessed by comparing the model values to a reference dataset (which should have been collected at three, or more, times the accuracy of the model being tested, according to ASPRS 1989). Building on the fact that the reference datasets provide more accurate representations of the bare earth than the filtered laser models, then the differences between them can be used as indicators of the location and amount of model error. These differences can then be analysed quantitatively in order to provide information about the characteristics of the errors. The simplest means of performing this analysis is through a qualitative investigation of the DTM, in which

the surface alone is observed for phenomena such as 'dimples' or 'depressions' which may show obvious errors. Although this method is quick and simple it provides no quantitative description and it is, of course, not always possible to detect errors within the DTM visually.

As an alternative, many researchers have used the differences between known and modelled height values at a given number of locations, and have used these to calculate the mean and standard deviation of the errors, and hence the RMSE (these are defined below). For example, the RMSE for USGS DEMs is calculated by comparing the DEM with 28 elevation values across the model surface (Wechsler 2003) - in a 30-metre USGS DEM, there are approximately 161,355 elevation values; the 28 verification points used for calculating the RMSE constitute 0.017% of the dataset. Similarly in the UK the Ordnance Survey's digital contour data have a quoted accuracy of +/- 1.0m in urban areas to +/- 1.8m in mountain and moorland regions.

The mean sample error is given by

$$m_\varepsilon = \frac{1}{n}\sum_{i=1}^{n}\varepsilon_i \tag{3.1}$$

where m_ε is the mean sample error, and ε_i is the error calculated at a given location defined as the difference between the reference dataset and the modelled value at that location. n is the number of points. The mean sample error will be zero only if the errors are unbiased. However, this measure does not provide any indication of the spread of values and as such the standard deviation should be calculated. This effectively provides a measure of variation within the errors (Zhang and Goodchild 2002).

The standard deviation of the error is given by

$$std_{\varepsilon} = \left(\frac{1}{n-1} \sum_{i=1}^{n} (\varepsilon_i - m_{\varepsilon})^2 \right)^{1/2} \tag{3.2}$$

In both theoretical and experimental analysis, the most frequent form of accuracy assessment is the root mean squared error (RMSE). The RMSE is a statistical measure of dispersion, and is defined as

$$RMSE_{\varepsilon} = \left(\frac{1}{n} \sum_{i=1}^{n} (\varepsilon_i^2) \right)^{1/2} \tag{3.3}$$

Clearly the larger the value of the computed RMSE, the larger the difference between the two datasets, in this case; the modelled elevation data and the observed terrain heights. Wood (1996) attributes the widespread use of the RMSE to the relative ease of this calculation and reporting (which is usually a single figure), and of course the ease with which the concept can be understood by most users of elevation data.

The limitations of the RMSE have been noted within the literature, including Carlisle (2000) who suggested that instead of using this measure we should create spatial distributed DEM error models, which he terms error surfaces, in order to better analyse DEM errors and uncertainty. Carlisle proposed that both the distribution and scale of errors within a DEM are related to the characteristics of the terrain. His assumption is corroborated by the findings of this investigation (see Chapters 6 and 7) in which the spatial pattern of error and apparent correlation with edge boundaries was noted. Carlisle (2000) proposes a technique in which differences between reference datasets and terrain models are used to compute DEM error - as in previous studies. Carlisle suggests that in addition to this a set of terrain parameters should be determined to characterise the terrain, and using these a regression

model should be developed in order to define the relationship between the DEM error and the terrain character. These regression models form the basis for creating an RMSE surface to portray DEM error. Clearly, these error surfaces provide more detailed information about DEM error than a single global error statistic such as the RMSE. According to Carlisle (2000) these error surfaces can be derived from known information about the relationship between error and terrain parameters. Terrain parameters include: elevation, gradient, plan and profile curvature, relative relief, texture, mean, minimum and maximum extremity, standard deviation and point distance. In his investigation Carlisle (2000) also calculated additional terrain parameters, resulting in a total of 96 parameters for each of the DEMs. Coefficients for the correlation between each terrain parameter and DEM error were calculated in order to provide an initial indication of the presence of a relationship. Multiple linear regression could then be used to define the relationship between the terrain parameter values and DEM error at 100 survey points. This equation can then be applied to the whole study area to produce an error surface. Clearly, the depth of this type of methodology is far beyond the remit of this investigation, not least because of practical limitations of time. Deriving 96 terrain parameters for each of the DTMs investigated here would have required a considerable amount of research time. It is worth noting that even had such an investigation been conducted the results may have actually have been of little generic value to DTM users: Carlisle (2000) found that there were differences in his results between different DTMs which somewhat throws into doubt the transferability of any conclusions derived from his investigation to other study areas.

Some of the objections to use of the RMSE relate to the statistics of the measure itself, while some relate to the way in which it is applied and, by implication, to the way in which it is used. The first problem, and the one which is of direct relevance to this thesis, is that the RMSE is a single, global descriptive value and as such cannot convey any spatial variation in the error across the DEM (this has previously been noted by a number of authors (Wood 1996; Monckton 1994). Further problems with the RMSE as a statistical procedure have been noted by Monckton (1994) and

Li (1988) Both authors advocate the reporting of standard deviation, and other measures (see below), in addition to the RMSE. As an alternative to relying on the RMSE calculations as a measure of DEM accuracy, Li (1988) suggests additionally using the range, or the two most extreme values of the differences between the DEM and the reference dataset. However, using the range may be misleading simply because it only relies on two values within the dataset and these are often outliers.

$$Range_{\varepsilon} = \varepsilon_{i(max)}) - \varepsilon_{i(min)} \tag{3.4}$$

Examples of the use of descriptive statistics for comparing DEMs are reasonably common, and include the work of Fischer and Dowman (2002) and Dowman and Fischer (2003) who used the RMSE, and other descriptive statistics of the calculated differences in their investigation of errors in IfSAR (Interferometric Synthetic Aperature radar) derived models. Li (1992) also advocated the use of simple descriptive statistics for modelling DEM accuracy, while Lopez (1997) based his analysis on the maxima of the differences between the reference and testing models. The use of descriptive statistics in this way is a valid technique for assessing global error, providing that the reference data are of higher accuracy than the model DTM. However, the RMSE does not provide any information about individual values within the DTM, such as how well each individual cell represents true elevation. There are alternatives to descriptive statistics which may be used to provide more information about the spatial variation of errors within a DEM.

In many respects a spatial statistical approach would seem to be the most sensible, particularly given that errors in spatial data are spatially autocorrelated (Ehlschlaeger 1998; Wechsler 2003). For example, an error in the value of one laser point will affect its neighbours during interpolation, and will consequently affect the form of the filtered DTM. This can create systematic errors in elevation models which clearly pose a problem for the use of non-spatial statistical methods to define accuracy, such as the RMSE, as these methods can not accommodate the spatial

autocorrelation of error. Given the problems of the global statistics approach to modelling error in elevation surfaces, a number of authors have advocated further studies, such as the provision of data quality information in the form of probabilities in order that the level of error may be compared against given thresholds (Zhang and Goodchild 2002). Others have promoted the mapping of uncertainty for discrete objects using epsilon error bands (Abeyta and Franklin 1998), whilst the mapping of the spatial pattern of error has been advocated by Foody and Atkinson (2002). A number of early studies, such as Guth (1992) offered preliminary investigations of the spatial variation of errors, showing that the errors identified in DEMs will tend to correlate with features in the topography. More rigorous analysis is now required to further explore the pattern of errors in elevation models.

There are other methods for modelling the spatial variation of DEM errors, although many of these would require a similar depth of investigation to that noted for the Carlisle (2000) method above. Such methods include that proposed by Pfeifer and Santos (2004) who employed the technique of least squares matching for DTM comparison. The result of their comparison is a vector field of the differences between two models. This vector field models the deviation from one model to another, with these deviations themselves being an indication of model quality. However, even using these methods there is a recurrent problem, as the results are communicated using global descriptive statistics. Moreover, the Pfeifer and Santos (2004) algorithm has only been tested using synthetic data.

Some researchers have used simulation techniques (such as Monte Carlo or conditional simulation) to generate a number of plausible DTMs from the dataset which could be compared to give an indication of uncertainty. Again, in terms of this investigation there were a number of practical problems with implementing simulation procedures mainly because laser research often involves using many millions of points which results in extremely lengthy computation times for each iteration. In addition, simulation can only really provide a model of uncertainty rather than of error, in other words it indicates where problems are likely to occur in any model but

it should not be used as an indicator of the magnitude of error. As an alternative to simulation, Brown and Bara (1994) used semivariograms and fractal dimensions to investigate the structure of systematic errors by analysing the anisotropy in DEM surfaces (i.e. where the general pattern of variation was different in one direction over a short distance). This anisotropy was interpreted as an indicator of systematic error, such as banding or striping. Polidori et al. (1991) also used fractal techniques to identify interpolation artefacts, and Lopez (1997) assessed the occurrence of random errors (which he found were weakly correlated with neighbouring points). Indeed, many of the spatial algorithms use the relationships between neighbouring points in order to understand error variation (e.g. Hannah 1981 and Felicísimo 1994). In effect, these methods compare the height values in a DEM to values obtained by interpolating from the neighbouring points. A threshold is set and the data point is considered erroneous when the difference exceeds this threshold. In such studies the variogram can be a particularly useful indicator of variation. For example, variogram modelling could perhaps be used to identify dominant forms of spatial variation for particular regions, following the example of Lloyd and Atkinson (1998). This study was used to describe variation in landform but could equally be used to describe variation of errors (which happen to be across landforms). In the method proposed by Lloyd and Atkinson (1998), variograms for moving windows are computed across the study area. Spatial variation of the phenomena under investigation is modelled using the form of these variograms. In a similar way, indices of spatial autocorrelation (such as Moran's I and Geary's C: Bailey and Gatrell 1995) could also be used to analyse spatial variation of error. Moran's I is closely related to the covariogram, whilst Geary's C is more related to the variogram. Analysis of the use of the variogram or covariogram for modelling the spatial structure of error in DTMs derived from laser data filtering would be interesting: however, it does not fall within the boundaries of this thesis. Despite the existence of a number of techniques for spatial modelling (as shown above) many of these employ methods which were not suitable for practicable inclusion within the context of this investigation. They are mentioned here in order to highlight potential avenues for extending the research.

3.5.2 Error in Continuous Models

Given that most elevation surfaces are represented as continuous fields, the accuracy of the field model in elevation modelling research is of major concern (Zhang and Goodchild 2002). Methods of accuracy assessment include Fourier analysis of topographic features (Balce 1987), mathematical modelling of DEM accuracy (Li 1994), theoretical modelling of accuracy (Li 1994), and uncertainty modelling in DEMs and their derivatives (Hunter and Goodchild 1997). As with errors in discrete models of elevation, errors in continuous surfaces may be concatenated and propagated in geographical analysis (Walsh et al. 1987) rendering their analysis crucial in any assessment of overall model accuracy. The accuracy of any continuous surface representation can be assessed objectively by calculating the mean and standard error from a set of samples that compare measured values against reference data (Zhang and Goodchild 2002). Accuracy may be interpreted and communicated in terms of normality of the data, as it can be assumed that the true value of the variable will not deviate more than twice the standard deviation from the mean value, with a probability of 95% Rees (2000). As with the modelling of error in discrete datasets, other authors have promoted the use of geostatistics for modelling the variation in accuracy of continuous surfaces. The theoretical and practical problems with this approach to urban modelling were introduced in the previous Section, and are discussed further in Chapter 6. In addition, the classification of categorical variables (which is a technique often used in remote sensing, including the modelling of ALS data) can be assessed by comparison against a reference dataset depicting land use. Most often, this technique involves using an error matrix in order to calculate the Percentage of Correctly Classified objects (PCC). Accuracies are compared using either the Kappa Coefficient or, more recently it has been argued that the z-test is a more appropriate method (Foody 2004). The merits of the z-test and a worked example are presented in Chapter 8.

In summary, the uncertainty in both objects and fields can be measured using statistics and misclassification analysis such as that detailed above. However, a more

important aspect of error description is the representation of spatial variation.

3.5.3 Significance of Errors in Elevation Models

To conclude this discussion of the methods of error modelling it is useful to examine why error assessment is important, and the ways in which the information might be used.

Several authors have assessed the significance of errors by investigating how error is propagated into products derived from the original model. Heuvelink (1998) conducted detailed studies on the modelling of propagated error. Heuvelink (1998) commented that despite the existence of methodologies for the modelling of error propagation, none of the techniques he investigated in 1998 had been incorporated within any professional GIS package. The importance of the requirement for error propagation information was also shown by Fisher (1990); (1991); (1992) who examined how errors in DEMs could affect the outcomes of viewshed calculations. The viewshed technique is used in numerous geographical applications, and is defined as the area which is observable from a viewing location versus that which is invisible (Fisher 1991). In his studies Fisher created an original DEM, and a number of test DEMs (which included error terms added to the original elevations) created using Monte Carlo simulation, to examine whether the viewshed area in the original was different from the simulated DEMs. It was found that the viewshed area in the original DEM was significantly different from those produced by the randomised process at the 90% significance level, where the viewshed in the original DEM was consistently greater in extent than that found in the same DEM with the addition of simulated error. Such conclusions have important implications for planning of, for example, visibility analyses such as those for military manoeuvres, and the planning for location of wind farms and landfill sites. Such investigations also demonstrate the inherent uncertainty of viewshed calculations, an observation which led Fisher (1991) to state that viewshed results should be treated with the utmost caution and

advocated the use of fuzzy viewsheds (as opposed to classic Boolean results, Fisher 1992) as a method of managing this uncertainty, as the fuzzy viewshed communicates some measure of uncertainty to the user.

Carter (1992) examined the effects of data precision on the calculation of slope and aspect. He investigated the cause of large spikes in frequency graphs of aspect angles calculated from DEMs which occurred in the cardinal directions. He determined that these spikes were caused by the technique of slope and aspect calculation, and specifically by the fact that elevations are often reported only to the nearest metre. It was shown in Carter's investigation that data precision had an important bearing on the ability to capture both aspect and slope values. Carter (1992) suggested that it is the responsibility of the data user to understand the limits of the data, and the errors which may exist within the dataset.

Li (1994) quantified the differences in accuracy caused by different DEM creation methods. He directly compared DEMs created from contour data, and those created from gridded data. In both cases he also quantified how the addition of feature specific data (such as points along breaklines) could reduce the error within the models. Other investigations regarding the management and reduction of error in processing of elevation models include Brown and Thaddeus (1994), who investigated the effect of systematic errors in DEMs, and Hannah (1981).

Of most relevance to this thesis are investigations which have examined the impact of data errors upon subsequent feature extraction from DEMs. Lee et al. (1992) quantified the effect of using different realisations of DEMs, which were created using Monte Carlo simulation, on the representation of a floodplain. They demonstrated that the changes in the magnitudes and spatial patterns of errors significantly affected the results of the extraction.

Despite the advances in understanding of errors in elevation models, as evidenced by the examples above, it is clear that most of work on error is still at the research level, although some more commercially oriented studies are now looking at the

conveyance and communication of error information to spatial data users (e.g Guptill and Morrison 1995). In particular the visualisation of error in DEMs has been the focus of much research, including Wood (1996) and Fisher (1992). Many authors have advocated the use of error information in commercial geographical analysis, with some suggesting that the onus should be on the data supplier to deliver quality evaluation along with the DEM itself (Monckton 1994; Foody and Atkinson 2002). For this reason, issues arising out of the communication of error are briefly touched upon in this thesis, and preliminary visualisations of identified error are produced.

3.6 Chapter Summary

Previous work from the related disciplines of urban surface modelling, urban remote sensing, and the measurement of error has been presented and reviewed. Particular focus has been devoted to a review of the techniques for modelling from ALS data, and the limitations of some of these techniques have been noted. What is apparent from this discussion is that the success of previous techniques has rarely been measured, resulting in the fact that many ALS data users are making uninformed decisions regarding the choice of post-processing methodology. Furthermore, they have a lack of tools within commercial packages with which to measure the success of their models. The implication of this has been highlighted in a discussion of the significance of errors in elevation models. Error assessment and quality evaluation of ALS data is important for several reasons. First, it may give important information about deficiencies of a post-processing approach and may thereby help to focus further research activities. Second, it is needed in order to compare the results of different approaches and to convince a user that an approach can be used in subsequent analysis for decision-making. For this purpose, evaluation should not only be based on visual and thereby subjective control, but on quantitative quality measures. The methods used to identify and visualise errors in ALS datasets are discussed in the following chapter.

Chapter 4

Data and Methodology

4.1 Chapter Overview

As a precursor to the following empirical research chapters (5-9), this chapter presents the research methods and datasets used within the investigation. The methodology presented is a framework of techniques which could be used for assessing the success of any ALS post-processing methodology, and includes methods for calculating, analysing and visualising error information. Methods for evaluating the quality of the methodology are also presented, in order to provide an indication of the validity and reliability of the research. As suggested in Chapter 1, the methods described for analysing errors are not restricted to use with ALS data, and could be used for the representation of other surfaces including other representations of urban form.

4.2 Aims of the methodology

The methodology was designed to answer the research problem identified in Chapter 1. To recap, this states that at each stage in ALS post-processing the user must

decide which algorithm, or combination of algorithms, is the most appropriate for their needs. In terms of error reduction, this requires that the user understands the characteristics of the errors which may be introduced by using any given algorithm, and chooses the method which reduces unwanted error. There has not been a focus within previous research on quantifying this error, and in many cases the user is currently unable to make an informed decision.

The aim of the methodology described here is to identify and analyse the error introduced at each of the three post-processing stages - data structuring, filtering, and segmentation - by using different methods and parameters for each process. In this way, each element of Equation (2.13) can be quantified and a full error budget be derived. The thesis underpinning the research is that the choice of post processing methodology will significantly affect the magnitude and spatial pattern of error within an ALS derived DEM. However, the research does not just examine whether there are significant differences in errors produced by using alternative post-processing methodologies, but rather seeks to determine the degree and the way in which different methods influence the characteristics of the errors produced. The error characteristics examined include magnitude of error, location and in some cases spatial pattern. From a user perspective the improvements offered by different post-processing algorithms are quantified. The findings of this investigation show the sensitivity to error of a variety of post-processing techniques. This highlights those areas where users need to make carefully informed decisions regarding their choice of methods in order to reduce error and maintain the high capture accuracies offered by this form of data acquisition.

The research tackled here has a wide scope. For clarification, this was divided into five separate investigations, each of which analysed errors introduced by different aspects of post-processing. The methodology for each investigation is presented in Section 4.5. In the investigations a variety of datasets were used, both for modelling and as reference datasets with which the models were compared and their relative successes compared. All of the datasets used within the research are described in

the following sections.

4.3 Data and Study Sites

Two airborne laser datasets were used for the four investigations. The first (a small area with lower point density) was used for the pilot study, the second was a larger area with a higher point density and was used for the main investigation. The parameters and conditions for data capture for each dataset are described here. A number of reference datasets were used in the investigations in order to quantify the errors in the post-processing workflow. Reference datasets included a high resolution photogrammetrically derived DTM, high resolution terrestrial laser points, an orthorectified aerial photograph, and a set of GPS readings of bare earth heights.

4.3.1 Airborne Laser Datasets

A pilot study was conducted in order to test the methodology for the main investigations. The term 'pilot study' tends to be used in two different ways within the literature. Firstly, it can refer to feasibility studies which are essentially small-scale versions (in the aspatial sense) and which are conducted in preparation for the main study (Polit et al. 2001). Secondly, a pilot study can be the pre-testing of a particular research instrument (Baker 1994), such as an algorithm or processing technique. In this instance, the pilot study was used to test for both elements - both to provide early indications of expected results, and to ensure that the designed tests and algorithms were acceptable. The aims of the pilot study were, firstly, to ascertain that the methodological design was practical for a much larger dataset and so to uncover any potential problems in the methodology. The second aim was to produce initial results which could subsequently be compared with the results from the main study in order to assess the transferability, or equivalence, of the results from this investigation. The suitable elements of the pilot investigations were subsequently

used in the main investigation, the aim of which was to ascertain the stability of the methods over larger areas, using higher point densities and data volumes. The main investigation also incorporated a wider variety of building types and urban land cover types.

The area chosen for the pilot project was a small area of Southampton, which was a convenient location for field work and testing. The airborne laser data for this region were supplied by the Environment Agency for England and Wales. The parameters for the capture of the pilot study data are shown in Table 4.1. The average point spacing was 2m, which was lower than that used for the main study: however this reduced point density (and associated reduction in dataset volume) was ideal for evaluating the methodology as processing times were correspondingly much faster.

Table 4.1: Parameters for capture of EA laser data

PARAMETER	SETTING
Aircraft	Cessna 404
LIDAR system	ALTM 2033
Flying speed	125kts
Flying altitude	850-900m AGL
GPS frequency	1hz
INS frequency	200hz
Laser repetition rate	33khz
Scan frequency	36hz
Scan width	+/- 13 degrees
Swath width	404m
Strip side-overlap	30%
Foot print size	23cm
Horizontal accuracy	+- 45cm
Vertical accuracy	+- 25cm

In order to test some aspects of the methodology, particularly for the restructuring investigation presented in Chapters 6 and 7, a subset of the data had to be used because of the lengthy processing times encountered when using large datasets. A case study of 1315 points was chosen which equated to an area of approximately 80m by 50m. This area was carefully selected in order that it represented a variety of urban structures. These included a large church, a flat roofed industrial building, and some areas of established vegetation. Importantly, it also included some areas

Figure 4.1: Orthorectified aerial photograph of the pilot study region showing area 80m by 50m. The church can be seen on the centre left of the image, distinct clumps of mature vegetation are clearly seen in the central region. On the right-hand side there are some industrial flat roofed buildings

of bare earth terrain within this region. The subset area chosen for the pilot study is shown in Figure 4.1.

The study area for the principal investigation was much larger than that of the pilot study and contained some 1.2 million laser points. The principal study site was a 1km^2 suburban area of Bristol, the location of which is shown in Figures 4.2 and 4.3. This area was chosen as it was considered to be representative of a typical urban area given that it contained a variety of residential houses and shops, some complex buildings such as schools, industrial buildings, and churches. The area also contained a variety of types of terrain typical of many urban areas, including fields, scrubland, and wooded areas and a range of bare earth surface types including grass and asphalt. This variety in the landscape was essential, as it was important that the filtering methods being compared were not biased to a particular type of urban terrain. The variety of urban landforms can be noted in Figure 4.4 which shows an aerial photograph of the study site. The terrain within the study region was fairly flat (with a general gradient of 1-3^o), however there was also an area of steep gradient which rose to a slope of 17^o in the lower right of the region. The change in

Figure 4.2: Map showing location of main study area in local context. Study area shown by circle. Scale of mapping 1:1 000 000. Reproduced from Ordnance Survey map data by permission of Ordnance Survey, © Crown copyright

gradient is shown in Figure 4.5 which depicts the contours in the region, and Figure 4.6 which shows a hillshaded image of the study area. The choice of the Bristol study site for the main investigation was also determined by the availability of data at the required resolution.

Laser data for the Bristol study site were supplied by Infoterra, and were captured from a fixed wing aircraft at a point density of 1.283 points per m^2 spatial resolution. The instrument used for capture was an Optech 2033 operating at a flying height of 850 metres, at a pulse rate of 33,000 pulses per second with a laser footprint of 230mm and a swath width of 619m. The processes swath width was 580m due to 1 degree trimming from each side of the swath. In accordance with such capture conditions the vertical accuracy in z was +/- 15cm (1 σ)(Andrew Anstee (Infoterra), Personal Communication, 2004). The data were acquired with a 39% lateral overlap

Figure 4.3: Map showing zoomed in location of the main study area, the boundary of which is shown by the solid red line. Scale of mapping 1: 25 000. Reproduced from Ordnance Survey map data by permission of Ordnance Survey, © Crown copyright

(which reduces to 35% when the 1 degree trim is taken into account). TerraMatch (part of the TerraSolid suite of software) was used to match the strips together by the data suppliers. TerraMatch examines overlapping strips of LiDAR data, by building a triangulated model of each strip and comparing overlapping strips with one other, and attempts to translate any observed differences into differences in Heading, Roll, Pitch and Elevation, which can then be compensated for. The GPS data were processed: the remote receiver on the aircraft and the master base station on the ground were combined to produce a GPS solution, which was then combined with the IMU data to produce a SBET (Smooth Best Estimate of Trajectory). This was subsequently combined with the laser data in REALM (Optech's custom data processing system) to produce the laser points. The laser points were then converted to OSGB36 using an OSTN02/OSGM02 based transformation.

Figure 4.4: Orthorectified aerial photograph of study site (1km^2)

Figure 4.5: Contours in the study region derived from a linear surface model (1m resolution)

4.4 Reference data

The reference datasets were used to quantify the errors within the post-processed laser data. These errors included inaccuracies in the predicted heights within the

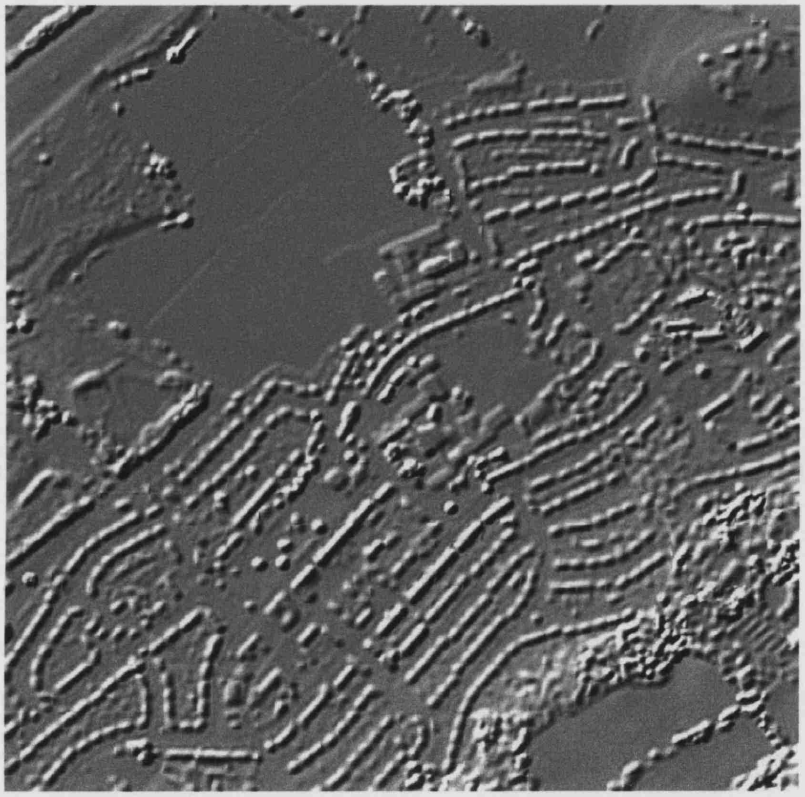

Figure 4.6: Shaded surface for 1m linear model derived from laser data

models, and misclassifications of the points or pixels. In order to ensure that the investigation was producing reliable results, a variety of reference datasets were used and the calculated errors from each were compared in order to ensure that consistent results were being obtained.

Three reference datasets were used in the main investigation to verify the success of the modelling from the ALS data. Each dataset was captured, or created, at a higher level of accuracy than the airborne laser data. Heuvelink (1998) suggests that reference datasets should be at least three times as accurate as the study dataset in order that they might be used to quantify error. In practice, this is difficult to determine. For the purposes of this investigation, it was considered that the improvements in positional accuracy (particularly in the z dimension) offered by the reference datasets ensured that they could be used in such a comparative way. The three reference datasets are described in this section.

4.4.1 Terrestrial Laser Data

In order to assess the success of the object modelling within the derived surfaces, and to estimate the amount of error in data capture, high resolution terrestrial laser data were captured for the study area using a Riegl LMS-Z420i class 1 laser. Terrestrial laser scanning effectively works on the same principle as airborne scanners working in either the green or the near infrared (NIR) parts of the spectrum. For the purposes of this study NIR was employed in order that a class 1 laser could be used. Laser safety is classified using a scale of 1 to 4 depending on the potential for causing biological damage. All lasers are labelled with one of these four class designations. Class 1 lasers cannot emit laser radiation at known hazard levels.

The LMS-Z420i instrument (shown in Figure 4.7) used in this investigation scans using a rotating mirror, which revolves continuously at an adjustable speed that was set at approximately 8000hz for the purposes of this investigation. The instrument has a range of up to 800m, which in practical terms meant that only a relatively

Figure 4.7: Dimensional drawing of the LMS-Z420i terrestrial laser scanning instrument. Diagram reproduced with kind permission of Dr Graham Hunter of Riegl

small number of scans (27) had to be taken to cover the study area. Each scan captured data from a 360^o sweep, with a vertical scanning angle of 80^o to allow for the capture of building roof morphology. The beam divergence for each point was approximately 0.25 radians. These settings rendered an effective point density of approximately 160 points/m^2 (although it should be noted that the exact point density depends upon the distance to the object being measured for terrestrial laser scanning). Measurement accuracy from the system is quoted by the manufacturers to be averaged at 0.5cm across the dataset, although Riegl technical specifications documents suggested that accuracies for individual points may be as high as 0.1 cm ($http://www.riegl.co.at$). A full description of contributing error sources for the terrestrial laser scanner is offered in Lichti and Gordon (2004), in which the importance of laser beam width and direct georeferencing are discussed.

In addition to the laser scanner itself, a digital camera and an RTK GPS kit were used during data collection. The mounting of these three pieces of equipment is shown in Figure 4.8. The camera used was a Nikon D100 with 14mm lens (this specification was chosen to ensure that each photograph covered the full vertical view of the scanner). Following capture of range data from the laser scanner, the Nikon camera took 7 photographs from a 360^o sweep.

Figure 4.8: Components of the equipment: at the base is the Riegl LMS-Z420i scanner, mounted on top is the Nikon camera, with the GPS antenna above

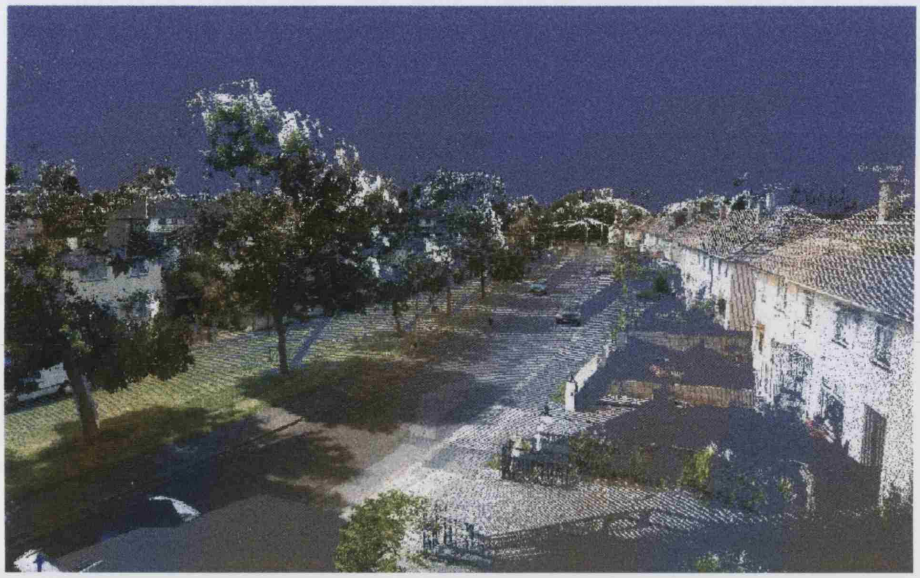

Figure 4.9: View of the terrestrial lidar dataset. Terrestrial laser points are coloured by RGB values attributed from the digital photograph captured at the same time as laser scanning

The RTK GPS kit used was a Leica System 500 SR530, which is a receiver based on Leica's ClearTrak technology. The instrument has a quoted accuracy of 2-4cm in the z dimension, however during capture in the field there were a number of sites where vegetation cover may have affected this accuracy. Accordingly, the recorded accuracy of each reading was noted during data capture. The accuracies ranged from 3.8 - 5.4cm, giving an average value of 4.5cm. This average value was used in calculating the overall error budget of the terrestrial laser capture in Chapter 5.

The recorded information from the laser scanner included range, angle and signal amplitude. The primary data provided by the LMS-Z420i were binary-coded point-cloud data in the sensor's own polar coordinate system. Further attributes to each measurement point such as intensity data or true colour (derived from the digital imagery) were also recorded, and attached as attributes to the raw laser data. The workflow for the matching and aligning of the individual scans (using information from calibration scans using retro reflective targets), and the amalgamation with the GPS data is described in Appendix C. To enhance visualisation of the data, the RGB values from the Nikon digital photographs were attached as attributes to the laser points following data capture. Figure 4.9 shows the coloured laser points from

one of the scans. The scan lines and pattern of data capture can be seen in these images.

4.4.2 Photogrammetric DTM

For comparison with filtered datasets, a ground reference model (DTM) was constructed, by photogrammetrists, from an aerial photograph using SOCET SET v4.4.2 terrain extraction and editing modules. The photography used in the DTM creation was captured at a low flying height (1700' above mean sea level (AMSL)) in order to gain a set of high resolution stereo imagery with a Ground Sample Distance (GSD) of approximately 5cm after scanning at 21 microns. Aerial triangulation was completed to a high accuracy with an RMSE for the block of 0.027m, 0.034m, 0.038m in x, y, and z respectively. The resulting imagery enabled the capture of details such as kerb offsets and step treads. A dense control network of 50 Ground Control Points (GCPs) was also used to enhance the DTM. The GCPs were captured using Real Time Kinematic (RTK) GPS to an accuracy of approximately 1-2cm in planimetry and 4cm in the vertical.

An initial TIN DEM was created by Mr Jon Horgan (Ordnance Survey) in which individual trees and buildings were removed using a set of parameters appropriate for the types of terrain present in the test area. The resultant TIN DEM was then edited with the Interactive Terrain Edit (ITE) module (in SOCET SET), enabling the inclusion of breakline data to enhance linear features such as road, rail and river networks. Isolated errors attributed to poor correlation spikes required individual editing in order to create a more accurate DTM. Where the terrain was steeper, a terracing effect became apparent, where the levelling of land affects the modelling of the terrain. Here many breaklines were required in order to accurately model the landscape. The TIN was then exported as a 1m grid for use in the comparisons. A number of control points were used in order to ascertain the accuracy of the reference model. The height values at these locations were determined using Real

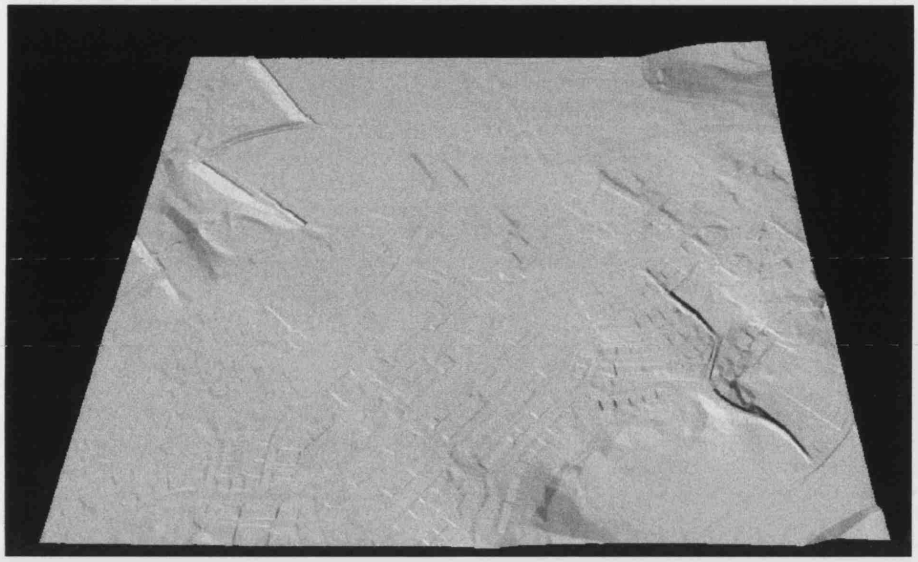

Figure 4.10: View of DTM created using photogrammetric techniques. Note the remaining footprints of many of the buildings, and the obvious smoothing of some areas

Time Kinematic (RTK) GPS by a team of field surveyors. The calculated accuracy of the photogrammetric model (RMSE) was approximately 11cm when compared to the independent GPS control. A view of the resultant DTM is shown in Figure 4.10.

4.4.3 RTK GPS Readings

To determine the accuracy of the filtered datasets created in Chapter 8 some additional RTK GPS readings were captured by field surveyors at a number of control point sites. The sampling strategy used for determining GCP location was restricted by certain accessibility constraints. Firstly, the location of the sites had to comply with survey health and safety regulations. Areas such as major roads and some industrial sites, whilst good candidates for flat bare earth terrain, are dangerous and so were deemed to be unsuitable for GPS surveying. In addition, there were some more obvious restrictions such as those points which were inaccessible, including many beneath buildings and in dense vegetation. Given these constraints it was not possible to use random, stratified sampling (which would have been optimal for

this investigation). It was considered that the best alternative was to take control measurements on areas identified as bare Earth in an orthorectified photograph.

Once captured, the points were processed using SKI-Pro software and the coordinates converted from the WGS84 datum to OSGB36 using the OSTN02/OSGM02 transformation and geoid model. The horizontal transformation accuracy of this procedure is 0.1m RMS whilst the vertical accuracy is 0.02m root mean squared error (RMSE).

4.5 Methodology

The datasets described above were used in the following methodology to analyse the characteristics of errors introduced during the post-processing of ALS data. This section is divided into four parts. The first details the methodology used for quantifying the errors within the ALS data prior to post-processing. This was a necessary precursor to the following investigations in which post-processing errors were quantified. The remaining three parts of this section present the methodologies used for analysing the error in data restructuring, filtering and segmentation.

4.5.1 Error in Data Capture and Preliminary Post Processing

The values in the raw ALS data were compared with those from the higher resolution terrestrial laser dataset. Vertical differences between the terrestrial and airborne range measurements were assumed to be, for the most part, attributable to errors in the airborne capture and preliminary post-processing. However, it was accepted that some differences would be attributable to temporal changes, as there was a ten month difference in the timing of data capture for the two datasets. The comparison was made on coincident points in the two datasets meeting a horizontal separation distance of 10cm. Measurements were taken on flat asphalt planes identified in the

datasets, such as road surfaces and flat roofs. These were chosen as those surfaces least likely to change between the capture of the two datasets. The differences between the two datasets were recorded, and descriptive statistics calculated. The results obtained from this investigation were compared with the quoted accuracy for the airborne instrument, and with results from previous investigations quantifying airborne laser capture accuracies.

The information resulting from this investigation was used in the remaining studies in order to quantify the errors introduced during subsequent stages of the post processing workflow.

4.5.2 Data Restructuring Investigation

In the first post-processing investigation, the errors introduced during the restructuring of data onto regular grids were analysed. There are two reasons why users may wish to restructure their ALS data. Firstly, many of the subsequent filtering and segmentation algorithms require a gridded input format, the reason for this being that many of these algorithms exploit pre-existing image processing techniques. Secondly, many software packages require that ALS data are gridded in order to display them and to calculate surface derivatives such as slope, gradient and aspect.

In order to restructure the irregular raw ALS data the user must decide how the 'new' gridded values should be calculated, and at what resolution this should be conducted. The calculation of the grid values involves using interpolation techniques. In this investigation, the effect of both the interpolation methods and the grid sizes were investigated in terms of the magnitude and spatial pattern of errors which were introduced. A pilot study was conducted to assess the suitability of the interpolation and spatial resolution parameters. In the pilot study (Chapter 6) five interpolation techniques and three grid resolutions were tested. The interpolation techniques included bilinear, bicubic, nearest neighbour, biharmonic spline and block kriging. The three grid spacings chosen were 1m, 2m and 4m so that the effect of using grids

both smaller and larger than the point density could be investigated. The effect of altering these parameters was assessed by comparing the interpolation values in the surfaces with the raw data values. It was assumed that any difference in these values could be attributable to errors introduced during the data gridding. There are many routines for spatial interpolation available and these have been widely documented in the past (e.g. Watson 1992; Burrough and McDonnell 1998). However, not all of the methods are suitable for urban elevation modelling from ALS data. In particular, for urban surface environments where there are frequent discontinuities, local interpolation rather than global or fitted function methods are preferable, in order to retain as much of the local complexity in the surface as possible. The description of the methodologies used in this investigation requires that the theory of interpolation be briefly introduced here.

Spatial interpolation may be defined as the procedure of estimating the value of a field variable at unsampled sites within the area covered by sample locations (Zhang and Goodchild 2002). The basic assumption underlying any interpolation procedure is that points that are close in space are more likely to be similar than points further apart. All interpolation algorithms therefore aim to estimate values at unsampled locations. However, there are many forms of spatial interpolation, which operate in very different ways, and which produce very different results. Watson (1992) states that the selection of the correct interpolation procedure is crucial to successful analysis of the data. The choice of the set of candidate algorithms was largely based on the suitability of the method for modelling height, and evidence from the literature. There are two principal methodological approaches for interpolation - deterministic and geostatistical interpolation, both are explored in this investigation. Deterministic interpolators use mathematical functions to drive the interpolation process, whilst geostatistical approaches use both mathematical and statistical functions for the interpolation.

There are different types of deterministic interpolation. Firstly, there are those which interpolate on the basis of similarities between neighbouring points - inverse distance

weighting is a widely used example (Longley et al. 2001). In these methods a small moving window is used to identify a set number of points. Such interpolation is used to provide a locally weighted average. In local interpolation only values which are within specified proximity to the unknown location are used for the computation. Such methods are generally used where there are no outliers in the data, and where the study area is evenly covered by data values. In this approach the influences of surrounding data points are computed and a set of weights/influences calculated for each individual point (this is called weighted average local interpolation). The fundamental advantage of this approach is that the surface produced is dominated by local trends in the data, and that it can cope well with large data volumes as each computation involves only moderately sized subsets of the data. The alternative type of local deterministic interpolation is the use of local polynomial functions. These methods are inexact, in other words the fitted surface does not necessarily pass through the sample points. The method produces a localised smooth surface whilst also preserving local detail. Such methods are suitable where the dataset exhibits short-range variations such as those found in urban environments.

The second type of deterministic interpolator is defined by the degree of smoothing in the resultant surface. Here all the points in a subset are used to derive a polynomial function which is then used to predict the values at unsampled locations: the interpolator is said to be global. These methods are also known as fitted function techniques or global polynomial functions. Global interpolators compute the surface from a polynomial function which is derived for the surface. The surface is fitted using least-squares regression. The complexity of the resultant surface is a function of the number of parameters, which is equal to the number of linear equations to be solved (Watson 1992). These methods are also inexact. In the fitted function approach, the interpolation method can be defined as a set of coefficients in a polynomial equation. The surface parameters are usually extracted by solving a set of linear equations which express the combined influences of the data and the criteria controlling the fit of the polynomial equation (Watson 1992). Fitted functions are therefore used to describe the effects of a particular value over distance or in other

words, the geographical variation over the surface. The product of a global polynomial interpolation is a smooth mathematical surface which represents the gradual trend in the surface over the study area. These methods tend to produce smoother surfaces simply because they impose a prescribed general behaviour on the surface to override noisy, or erroneous data. However it has been noted (Watson 1992) that fitted functions have a tendency to overshoot in situations where a tighter, sharper curve established by weighted averages would suggest more conservative changes in relief. These methods are most suitable where there are no outliers in the data, and where the surface varies gradually over the region. Clearly, there may be problems with using such functions in dense urban areas, which are characterised by discontinuities and uneven change in height. With these methods, a certain degree of local detail is submerged in the overall surface. In summary, weighted average methods emphasise local detail, whereas fitted functions summarise global behaviour. It is also important to note that there are very different computational requirements for the two methods, namely that the weighted average method requires considerably more time per interpolation point than do fitted function methods.

The alternative to the deterministic approach is to use geostatistical methods. These techniques are founded on statistical concepts, which represent the arithmetical relationships between raw data. In addition to surface prediction, geostatistical methods provide a measure of the accuracy of the prediction (in the form of a variance function). Geostatistical methods, known as kriging, are more complex than the deterministic approach as they require an understanding of the statistical principles of spatial autocorrelation. However, this complexity is justified for interpolation where the variation of an attribute is very irregular, and the density of observations is such that the simple methods of interpolation might yield unsatisfactory results (Burrough and McDonnell 1998). Kriging is an interpolation technique which models spatial variation across a surface by looking at general trends/patterns and also by modelling the residuals from this (ie. the fluctuations which are not explained by the general trend). In this way kriging is more robust than other methods, which can tend to oversimplify the spatial pattern of variables, and it is often claimed that

kriging provides the most accurate predictions of all interpolation procedures (Armstrong 1998). A variety of commercial software packages offer simple geostatistical methods, although Burrough and McDonnell (1998) consider that it is more satisfactory to export the data to a specialised geostatistical package. This recommendation is adopted in this study.

There are many different types of kriging, not all of which are suitable for terrain modelling. Ordinary kriging (OK) was thought here to be the most appropriate for the modelling of terrain data. OK requires the local mean for each location to be estimated. Whilst it is noted here that kriging was originally designed to map the variation of naturally occurring phenomena (e.g. mineral deposits), the fact that this method accommodates both local and global trends may make it more suitable than other methods for modelling height surfaces in the built environment. The methodology employed in this investigation was designed to test this.

In this study one geostatistical, and four local deterministic methods were chosen for creating regular gridded data from scattered input data. Only local deterministic interpolators were chosen as it was considered that the global approach would be inadequate for modelling the discontinuities in the urban environment. The methods chosen were those most commonly used, namely: nearest neighbour, bilinear, bicubic, biharmonic splining, and ordinary block kriging (OK). The techniques for the bilinear, bicubic, and the nearest neighbour are documented in Watson (1992), these methods are all based on a Delaunay triangulation of the data. The method for the biharmonic spline is shown in Sandwell (1987). The deterministic interpolation techniques were implemented using MatLab (v7) software. Two packages were used for geostatistical interpolation: Gstat (described in Edzer and Wessling 1998) was used for the computation of the semivariograms, whilst the kriging was conducted using the GsLib programme (Deutsch and Journel 1992). Details of all of these interpolation methods are presented below.

4.5.2.1 Nearest neighbour interpolation

This method is the simplest of all the methods investigated as it does not require any numerical calculations. The method fits a piecewise constant surface through the data values, where the value of an interpolated point is the value of the nearest point. As the nearest neighbour method is an exact interpolator: no new data values are computed, and the value of the nearest raw data point is simply transferred to the unknown sample location. This method generally performs poorly for numerical data, particularly where these fluctuate spatially. Furthermore, the method cannot model oblique slopes: instead it produces a stepped surface, which preserves any discontinuities in the original surface, but also creates new false breaklines (Zinger et al. 2002).

4.5.2.2 Bilinear interpolation

In brief, this method fits a linear surface through existing data points in two dimensions. Two iterations of interpolation are conducted in this study. In the first the irregularly spaced raw data are interpolated onto a regular grid, and in the second stage the regularly spaced interpolated points are used to interpolate back to a retained set of raw data locations (in order to assess the accuracy and stability of the method). These two approaches require different bilinear techniques. In the first iteration, the irregularly spaced data are triangulated using Delaunay triangulation. In this approach the longest side of any triangle is minimised, thus reducing the distance over which interpolation must occur.

The three values at the nodes of each triangle are then used to model a plane. When viewed in three dimensions, each triangle essentially forms a facet of the surface, and together the facets approximate a continuous surface model. Based on the three corner points, each planar triangular facet can be fitted:

$$Z = a + bx + cy \tag{4.1}$$

Where a, b, and c are fitted coefficients from the three nodes of the triangle.

This linear surface model can be used to interpolate elevation values for an unknown location (x, y) within any facet. This interpolation process is a weighted average, and takes into account both the values at the three vertices, and the distance between the vertices and the unsampled location. The elevation of the interpolated surface Z at (x, y) within the surface is given by:

$$Z(x, y) = \Sigma_{i=1}^{3} w_i f(x_i, y_i) \tag{4.2}$$

Where: w_i is the i^{th} barycentric coordinate (see Watson 1992, pg76) of the interpolation point with respect to the triangle, and $f(x_i, y_i)$ is the observed dependent variable (the height) at the data point (x_i, y_i). As the sum of the barycentric coordinates is always one, each interpolated value lies on a plane that is usually tilted and coincides with the three data values at the corners of the triangle. Although the linear planar fitting on the triangles produces a continuous faceted surface, the surface slope changes abruptly when crossing over edges between adjacent triangles.

In the second iteration, the values at irregularly spaced locations are interpolated from a regular grid. This requires a different methodology, where the value of an interpolated point is a combination of the values of the four closest points. The example shown in Figure 4.11 shows how the method works in practice. The value at unknown sample site $(\hat{Z})$ is calculated from the four surrounding points $(Z_1 - Z_4)$. In the first stage of the interpolation the value at Z_i is calculated by linearly interpolating between Z_3 and Z_4, and separately Z_{ii} is calculated by interpolating between Z_1 and Z_2. Any point, along the lines connecting these pairs of respective points, can then be predicted using the straight line equation describing the slope

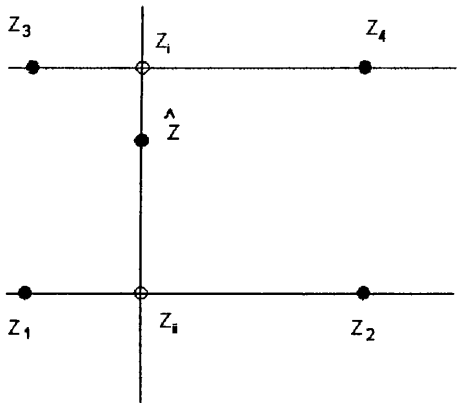

Figure 4.11: Example of bilinear interpolation

of the line. The straight line through the unsampled point $\hat{Z}$, which intersects these two straight lines at any two points, here Zi and Zii, can then be used to interpolate the value at $\hat{Z}$ as the values at Zi and Zii are already known. Bilinear interpolation thus takes the form of a polynomial equation, involving powers of x and y no greater than 1. It has four coefficients: $Z = a+b*x+c*y+d*x*y$. This method is piecewise bilinear, and is faster and less memory-intensive than bicubic interpolation. It is up to five times slower to compute than a nearest neighbour interpolator on the same data. Bilinear interpolation will produce a smoother surface than the nearest neighbour technique, and will preserve local effects in surface change.

4.5.2.3 Bicubic interpolation

For the first stage of interpolation, the method fits a bicubic surface through existing data points which have been triangulated. Bicubic interpolation is continuous in the first derivative, which means that cubic surfaces are often smooth, as the gradients between the piecewise segments are continuous. Bicubic interpolation for regularly spaced data, often termed cubic convolution, works in a similar way to the bilinear method but for a regular grid uses the sixteen surrounding points in interpolation. The value of an interpolated point is a combination of the values of the sixteen closest points. The method used for this investigation is a piecewise bicubic polynomial, and produces a much smoother surface than bilinear interpolation. It tends to pro-

duce a smoother resultant surface than either the nearest neighbour or the bilinear interpolators. It also represents the statistical properties of the original data simply because it does not merely smooth out the values. Bicubic interpolation is recommended when the interpolated data and their derivative must be continuous. It produces a smoother surface than bilinear interpolation, but it takes approximately seven times longer to compute than a bilinear interpolation of the same data.

4.5.2.4 Biharmonic Splining

There are a number of different types of spline which are generally differentiated from one another by the number of constraints placed upon the function. This is determined by the order (m) of the polynomial used to define the particular spline. The number of derivatives which constrain the spline is given by $m - 1$, such that a quadratic spline must have one continuous derivative, whilst a cubic spline must have two continuous derivatives at each knot (a knot is a point on the curve). The term bicubic spline is used for the three-dimensional situation in which surfaces instead of lines need to be interpolated. Generally, cubic splines are more complicated than basic cubic interpolators and tend to produce a smoother surface for the reason that cubic splines have two continuous derivatives (Burrough and McDonnell 1998). In other words, continuity is ensured in:

- Elevation (zero-order continuity): the surface has no cliffs
- Slope (first-order continuity): the slopes do not change abruptly
- Curvature (second-order continuity): minimum curvature is achieved

The basic cubic spline connects the given data points by third order polynomials over each interval. Rapid changes in the gradient may occur in the vicinity of the data points, and the main draw back of the basic minimum curvature spline is the generation of overshoots in the regions with rapid change of gradient. Watson (1992)

noted that this was particularly problematic where the range of heights is large, and if data are scattered. Both of these factors are applicable to laser scanning data over urban areas. Watson (1992) also noted that there may also be distortions of the splined surface at the edge of the dataset. This effect is investigated as part of this thesis. There are further varieties of splines which impose further minimisation criteria upon the function. The biharmonic spline used in this investigation is a minimum curvature spline. It differs from the basic spline in that it imposes minimisation criteria upon the function. In the biharmonic spline, the square of the Laplacian is minimised which means that this spline has zero tension (this is explained further below). Although minimum curvature splines produce a visually pleasing smooth surface, in areas of large changes in the surface gradient the method can create strong artificial oscillations in the unconstrained regions. In other words, the method may produce large overshoots in the interpolated surface. In order to avoid such overshoot lower-order methods may be employed (such as minimising the power of the gradient). However this also removes the smoothness constraint and can lead to gradient discontinuities. The classic remedy for this problem is to use splines in tension, which are constructed by minimising a modified quadratic form that includes a tension term. As stated above, the biharmonic spline has zero tension, the opposite of which would be infinite tension. In the case of infinite tension the interpolation surface would be constrained such that its local extrema only occur at the input data locations. Tension splines behave similarly to cubic splines, with the addition of the tension parameter which just specifies how tightly the spline forms to each knot. A low tension parameter will yield a spline similar to a cubic spline; whereas at high tension, the spline will approach a linear surface form.

In summary, in the output from a biharmonic spline the maxima and minima do not necessarily occur at the data points. It is a local interpolator and can be exact, or can be used to smooth a surface. The computing load is moderate although slightly more than the bicubic method. Splines generally produce poor results where the surfaces show a marked fluctuation, as this can cause severe oscillations in the spline. However, splined surfaces are aesthetically pleasing and quickly produce an

overview of the data, and because of the smoothness of the surface the mathematical derivatives are easily calculated for the purpose of further analysis. Splines do not offer a direct estimate of the errors they introduce. These errors may be significant because of the nature of the unrealistically smooth surface which is created. For subsequent feature extraction this smoothing may generate misleading results.

4.5.2.5 Geostatistical Interpolation: Kriging

> Of all the common methods of spatial interpolation it is kriging that makes the most convincing claim to be grounded in good theoretical principles. Longley *et al* (2001)

Kriging interpolation is part of a family of analytical tools called geostatistics and was developed in light of observations that in some instances, the spatial variation of continuous attributes is too complex to be modelled by a smooth mathematical function. In the kriging approach, it is assumed that the complex spatial variation of a phenomenon can be split into three component parts: firstly, the general pattern of spatial variation - this is a deterministic function which describes the structural component of the variable (z). This function is denoted by $m(x)$. Secondly, the spatial variation is also explained by the residuals from the deterministic function $m(x)$. These are random fluctuations which vary from the overall pattern described by the deterministic function. There are two types of random component, the spatially dependent residual and a spatially independent residual. The spatially dependent residual (e') varies locally but generally follows the pattern defined by the deterministic component $m(x)$. The spatially independent residual (e'') is a spatially independent Gaussian noise term. These three terms $m(x)$, $e'(x)$ and e'' together explain the spatial variation of the variable Z at any location (x), so that:

$$Z(x) = m(x) + e'(x) + e'' \tag{4.3}$$

In the simplest case there is no trend or drift ($m(x)$), in other words given an entire population from which a number of samples might be taken, there is no difference in the mean values of the samples. In this respect there is a constant mean. For example, say that the IQ of all 15 year olds in four hundred schools nationwide is measured. Given the mean ($\overline{x}$) of each sample, we would not expect there to be a difference between the average IQ of pupil samples across the country. In this case the mean is constant and there is no trend in the population. The deterministic function therefore equals the mean, and therefore the expected difference between any two places x and $x + h$ (separated by vector h) will be zero:

$$E[Z(x) - Z(x + h)] = 0 \tag{4.4}$$

Where $Z(x)$ and $Z(x + h)$ are random variables at locations (x) and ($x + h$) (after Burrough and McDonnell 1998)

Given that the mean is constant, it can be assumed that the difference between the value of two variables is a function of the distance between them, The expected value at any location $Z(x)$ is therefore defined by the random spatially dependent function, the value of which depends on the distance between sites:

$$\begin{aligned} &E[(Z(x) - Z(x + h))^2] \\ &= E[(e'(x) - e'(x + h)^2)] = 2\gamma(h) \end{aligned} \tag{4.5}$$

Where $\gamma(h)$ is known as the semivariance (after Burrough and McDonnell (1998)).

In other words, the value of Z at (x) is not a function of the trend in this case, because it has been established that the mean is constant, instead it is explained in terms of variations in the random variable e'. This random variable is called

the regionalised variable, and is central to geostatistical analysis. What it means is that once the structural effects have been accounted for, any remaining variation between sites is homogeneous in its variation, so that the differences between sites are merely a function of the distance between them. The kriging approach thus describes the rate at which the variance between points changes over space. This relationship is formally modelled in the semi-variogram, which shows how the average difference between points changes with distance. If the conditions of stationarity (no trend/constant mean) and the variance of difference (differences depend only on the distance between sites) are fulfilled then the sample data can be used in order to estimate the semivariance using the following equation:

$$\hat{\gamma}(h) = \frac{1}{2n}\Sigma_{i=1}^{n}(z(x_i) - z(x_i + h))^2 \tag{4.6}$$

Where: n is the number of pairs of sample points of observations of the variable z separated by a distance of h known as the *lag*.

The plot of semivariance ($\gamma(h)$) against h (the experimental variogram) represents the first stage towards a quantitative description of the regionalised variable. A theoretical model is then fitted to the points in the experimental variogram, and the results of this fed into a parameter file for the subsequent kriging.

There are many variants of kriging, and the choice of which to use in any particular investigation depends largely on the results from a preliminary analysis of the raw data. The preliminary investigation involves examination of a histogram of the height values for a particular study. In the case of this investigation, the distribution of values was strongly negatively skewed. In such cases, it is generally noted (Burrough and McDonnell 1998) that three choices are available for subsequent kriging:

- Create a log-normal distribution, and use log-normal kriging to model the surface

- Krige in the presence of a trend (universal kriging (KT)), in this case some attempt is made to model the trend
- Ignore the skew and use ordinary kriging

Each of the above options were investigated to assess which was the most appropriate for the pilot investigation. Log-normal kriging is the interpolation of log-normally distributed data (Deutsch and Journel 1992). The predictions resulting from the kriging must then be back-transformed using an anti-logarithm. The extreme sensitivity of errors arising from antilog back transformation has been previously noted (Deutsch and Journel 1992), this renders lognormal kriging very difficult to use in practical research situations. The study data for this investigation were logged in accordance with the Deutsch and Journel (1992) methodology, however this did not return a normal distribution. Given the known problems associated with this technique it was decided that log-normal kriging (with a non-normal transformed data set) would be unlikely to yield more accurate results than ordinary kriging. In such cases Deutsch and Journel (1992) advocate the use of multi-Gaussian (MG) kriging, although this is a rare form of kriging and there is little published about its theory or practical application, and as such it was discounted for use within this study. The second option was to use universal kriging (KT), in which the knowledge from empirical (regression) transfer models is incorporated into the kriging equations. However, this form of kriging has tended to be used by environmental scientists and other modellers of natural phenomena and its transferability to modelling urban surfaces for this investigation was not known.

Given the above, the decision was made to conduct the investigation based on ordinary kriging (OK) and ignore the skew as recommended by Burrough and McDonnell (1998). OK assumes that the surface variation is statistically stationary and free of any systematic trend. In this type of kriging the weights for interpolation are chosen so that the estimate of the variable is unbiased and the variance of the estimate is less than for any other linear combination of observed values. This type of interpolation is exact, in other words the interpolated values will coincide with the

values at the original data points. Block kriging is a type of OK in which the variogram is used to estimate areas (termed blocks). This was thought to be the most suitable form of kriging for use in this investigation in the restructuring of ALS data onto a grid (P M Atkinson, Personal Communication 2003). Block kriging has also tended to be used to eliminate spikes and pits in surfaces of natural phenomena, and was also considered suitable given the known short range variation of elevations in urban environments. For block kriging the original kriging equations are modified to estimate average values of z over a given area (in this instance a cell). When block kriging a discretisation block is placed around each interpolation point. The interpolation estimate for the point is based on the mean value of estimates for each of the discretisation grid points. Estimation variances for block kriging are quoted as being lower than for point kriging (Burrough and McDonnell 1998).

Notably, inverse distance weighting (IDW) was not chosen for this study despite the fact that it has been used in a number of similar studies (Lloyd and Atkinson 2000). Whilst IDW is one of the most common interpolation methods used in many GIS analyses, it not best suited to the creation of terrain surfaces because of the nature of the principal underlying assumption that the variable being interpolated decreases in influence with increased distance from the sample point locations. This assumption is not valid for terrain data, and surfaces created using IDW tend to have a undesirable dimpled effect at the sample point locations (Longley et al. 2001).

The MatLab code used for the restructuring, and analysis of, the laser data is presented in Appendix B. These scripts were all written as part of the Ph.D research, except where otherwise stated.

4.5.3 Quantifying and Analysing the Error in Re-structuring

In the pilot study each of the methods described above was used to model the surfaces at different spatial resolutions (1m, 2m, and 4m). The error introduced at each spatial resolution was quantified and mapped. The principal objective of this

investigation was to quantify the amount and distribution of error introduced by interpolation. This was assessed by gauging the success of the surface modelling. Clearly for the deterministic methods, if the data points are used to interpolate a surface, then the goodness-of-fit of the surface cannot be assessed with these same data points as this would yield an overly optimistic (low) prediction error (Martinez and Martinez 2002). For this reason the accuracy of each interpolation method was assessed using standard validation procedures which involve omitting some raw data from the interpolation procedure and assessing the success of the procedure to model in the absence of these known values. There are a number of techniques which employ this principle, including split-sample testing, cross validation and jack-knifing. Whilst these three techniques are standard methods used in statistical and applied industrial mathematics research (for example, Efron 1982; Efron 1993; Tomczak 1998; Martinez and Martinez 2002), they have not previously been used for assessing the spatial pattern of errors in ALS models. The three techniques were used to reduce the risk associated with relying on the results from one test only, and this accordingly increased the reliability of the investigation as equivalence was tested for between the results from all three tests. Three techniques are explored below.

4.5.3.1 Split-sample Testing

Declercq (1996) advocates the use of the split-sample routine which can be used to assess the stability of the interpolation algorithm. In other words, this technique can be used to assess how the methods perform given a diminishing number of data points. In this procedure some of the sample values are omitted, the interpolation is performed, and the difference between the predicted and the raw data values at each omitted sample location is calculated. This difference is then used as a measure of the success of the algorithm. This method is also referred to as true validation (Voltz and Webster 1990), and has been used previously by Desmet (1997) in a study looking at the effects of interpolation errors on bare earth DEMs. The

usefulness of this technique is increased when it is used iteratively, where the number of omitted points is increased progressively and the differences calculated. This can return useful information regarding the stability of the algorithm, and its ability to cope with differences in input point density. For ALS modelling it is useful to understand how differences in point densities can affect the surface model. Such density variations may arise during data capture where there may be a lower density or absence of points over for example a snow patch on the ground or a glass roof of a building where there may be no reflection recorded.

The split-sample method was used by Lloyd and Atkinson (2002) using sample sizes of 95%, 50% and 25% of the original points to measure the effects of interpolation for rural DEMs. The methodology outlined in Lloyd and Atkinson (2002) is adhered to in this study, and the same proportions of omitted data are used here for the comparison of interpolation methods. For the grid spacing investigation a random selection of 5% of the raw points was omitted and the surfaces produced at different spatial resolutions. The number of omitted data points in each sample was not varied for this part of the investigation.

In the split-sample investigation the surface models were computed using 95%, 50% and 25% of the raw points, and the resultant surface model used to predict the values of the omitted points. The differences between the predicted and the omitted values (model error) were calculated for each surface at each spatial resolution, where

$$e(x) = z(x) - z^*(x) \tag{4.7}$$

where $z(x)$ is the data value at location x and $z^*(x)$ is the estimated value at the same location

Both the deterministic and the geostatistical methods were assessed using the split-sample routine as this method indicates the stability of the algorithms given progressively fewer input data points. In each investigation the omitted data points

were chosen randomly, and the tests run three times to check for consistency.

Whilst split-sample validation can give an indication of the amounts and patterns of errors created by interpolation algorithms, it cannot be used for quantifying the model error with any certainty as the computed model error will differ depending on which points are randomly chosen for omission. Moreover, because more than one point is omitted at a time the calculated model error is not a true reflection of the amount of error introduced by a particular scale and method. For this reason, a number of complementary tests were also run as part of this investigation.

4.5.3.2 Cross validation

Cross validation is a procedure somewhat similar to the split-sample method in that both methods partition the dataset, and both use one set to predict the surface, and the other to test this surface. However, cross-validation differs from the split-sample routine in that it is an iterative process. It is most often used as an exploratory technique to find the most suitable model among a number of candidates (Davis 1987; Deutsch and Journel 1992), and it has also been used previously in comparisons of interpolation algorithms (Tomczak 1998). The most common form of cross-validation is the leave-one-out technique. This involves removing one point at a time from the dataset, performing the interpolation, using the interpolated surface to predict the value of the removed point, and calculating the difference between the removed data point and the actual value of the removed point. In this form of cross-validation only one value is removed at a time, and the whole process is repeated until every sample point has been removed in turn. The overall performance of the interpolator is then evaluated as the root-mean of squared residuals (Davis 1987; Martinez and Martinez 2002). Low RMSE indicates an interpolator which is likely to yield reliable estimates for unsampled regions.

However, whilst cross-validation may yield more reliable information than the split-sample method, it is an exploratory rather than a hypothesis testing technique as

it cannot confirm that a particular model is, or is not, the optimum. A number of authors, including Davis (1987), warn against using the results in a confirmatory way, but do advocate using it in conjunction with the jackknife technique.

4.5.3.3 Jackknifing

The jackknife is another data partitioning routine which has some similarities to the cross-validation method described above. The jackknife is used to estimate the bias and the standard error of statistics (Efron 1982; Efron 1993; Martinez and Martinez 2002). Originally the jackknife technique was used to reduce bias (Quenouille 1956; Davis 1987). However, it is used in this investigation to draw conclusions about an estimator's sensitivity to individual observations and thus offer an indication of the stability of the interpolation method. The jackknife method may be summarised as follows (based on Martinez and Martinez 2002, and Tomczak 1998):

- Remove one observation from the sample data (the ith point)

$$z_1, ..., z_{i-1}, z_{i+1}, ..., z_n \tag{4.8}$$

- Calculate the value of the statistic of interest using remaining sample points to obtain $T^{(-i)}$

$$T^{(-i)} = t(z_1, ..., z_{i-1}, z_{i+1}, ..., z_n) \tag{4.9}$$

- Repeat until all values have been omitted in this way
- Calculate the jackknife estimate of the bias ($Bias_{Jack}$) for each interpolation method at each spatial resolution
- Calculate the jackknife estimate of the standard error ($\hat{SE}_{Jack}(T)$) for all of above

The estimate of the jackknife statistic (T) (which is calculated using the reduced jackknife sample) is given by

$$\hat{Bias}_{Jack}(T) = (n-1)(\overline{T^{(J)}} - T) \tag{4.10}$$

where

$$\overline{T^{(J)}} = \frac{\sum_{i=1}^{n} T^{(-i)}}{n} \tag{4.11}$$

Note that $\overline{T^{(J)}}$ is simply the average of the jackknife replications of T.

The estimate of standard error for each surface can then be calculated using the jackknife as:

$$\hat{SE}_{Jack}(T) = \left[\frac{n-1}{n}\sum_{i=1}^{n}(T^{(-i)} - \overline{T^{(J)}})^2\right]^{1/2} \tag{4.12}$$

In summary, three validation techniques are used in this investigation to compare the errors (that is, their magnitude and spatial patterning) created by interpolating at different spatial resolutions, with different methods across an urban surface. The results from each validation method reveal different characteristics of the methods and the effect of spatial resolution.

4.5.3.4 Validation of the Kriged Surfaces

Evaluation of the Kriged surface does not require recourse to cross validation and jackknifing routines. Both of these validation routines require extra validation points that were not included in the original interpolation. One of the benefits of the geostatistical approach is that it provides both interpolated/predicted values *and* an estimate of the uncertainty associated with this value. The validation of the

Kriged surfaces for the different spatial resolutions is therefore assessed using the estimate, or variance, values computed during kriging. These may be mapped in a similar way to the deterministic errors in order to assess the spatial pattern of errors.

These validation techniques were used for the pilot study to assess which of the interpolation techniques were most suitable for the main study. On the basis of the pilot study results (which are presented in Chapter 6), the main study involved a comparison of three interpolation methods at three spatial resolutions. The interpolation methods chosen were bilinear, nearest neighbour, and bicubic at 0.5m, 1m, and 2m grid resolutions. These grid spacings represented resolutions both smaller and larger than the original point spacing. The success of the modelling for the main study area was assessed using three split-sample routines. Cross-validation and jack-knifing were not practical methods for use with the main study area given the number of laser points. Without additional computational resources, it was estimated that using cross-validation for the study area in this investigation would take in the region of three hundred days to compute using a SUN sparc ultra5 unix machine running the Solaris operating system (version 8), with 2Gb swap space.

The success of the surface modelling for the main study area was further assessed by comparing the heights within the model with a sample of terrestrial laser data. The differences between the two are calculated and analysed. The patterns of errors are commented upon and compared with ground classifications to ascertain whether certain errors are associated with particular land classes.

The data restructuring investigation described here was designed to test the hypothesis that different data gridding methodologies produce significantly different errors in the predictions of heights in the resultant surfaces. If the results from the test support this assertion then this should be of interest to laser data users, as the results will firstly demonstrate the importance of choosing an appropriate gridding methodology, and secondly will indicate which parameters should be used to obtain

desired results, such as the reduction of error or the preservation of edges. The investigation and the subsequent results are described in Chapters 6 and 7.

4.5.4 Data Filtering Investigation

The second investigation focused on quantifying the error introduced in the filtering stage of data post-processing. Filtering is the process of classifying laser points (or pixels) into either bare earth or above ground classes in order to produce filtered datasets containing ground points only. Errors may be introduced during the classification process or in the subsequent modelling of the bare earth to form a continuous surface. It was considered that any errors greater than those identified for data capture would be attributable to errors introduced by filtering parameters.

The investigation was designed to assess how each of the following factors affected the amount and location of error in filtered datasets

- input data structure
- spatial resolution
- choice of filtering algorithm
- filtering parameters

To assess the affect of these factors, three filtering algorithms, each with two different settings, were used to filter both raw and gridded data (the gridded datasets were at three different resolutions as in the previous investigation). The resultant filtered datasets were then compared with reference data in order to ascertain whether there were significant differences, and if so where these differences occurred.

Of the three filtering methods used in this investigation, two were designed for use with raw, irregular data, and one for gridded data. The methods chosen were Iterative Robust Interpolation (IRI), Adaptive TIN Models (ATM), and geometric

filtering. The two filtering methods designed for irregular data (IRI and ATM) use a discrete data model, however gridded data can also be filtered using these methods if the values of the grid nodes are taken as discrete input points. In this way the result of data restructuring for the filtering stage can be analysed. The filtering methodologies, and variations in the parameters used, are described in detail in Chapter 8. The choice of filtering methods was based on the results from previously published comparisons such as that of Sithole and Vosselman (2003).

The utility of each filtered dataset was compared in terms of classification accuracy and height modelling accuracy. To assess the utility of the classification, contingency tables were calculated for each dataset. The contingency tables were then compared to assess if they were statistically significantly different from each other. Many previous studies which have investigated the statistical significance of differences in classification accuracy have based the analysis on the comparison of accuracy expressed in terms of the kappa coefficient of agreement. However there are a number of concerns which have been presented regarding the use of this measure (e.g. Foody 2004; Stehman 1997). Given the contentiousness of using the kappa coefficient it may be preferable to use alternative metrics for accuracy assessment, such as those presented in Foody (2004) one of which is the z-test for independent samples.

The z-test is a test for the significance of the difference between two proportions, and is estimated from

$$z = \frac{\frac{x_1}{n_1} - \frac{x_2}{n_2}}{\sqrt{p(1-p)(\frac{1}{n_1} + \frac{1}{n_2})}} \tag{4.13}$$

where x_1 and x_2 represent the number of correctly allocated cases in two independent samples of size n_1 and n_2 respectively, and $p = (x_1 + x_2)/(n_1 + n_2)$.

The statistical difference in accuracy between the two classifications was evaluated through a comparison of the calculated value of z (calc-z) with the tabulated value

(tab-z). Eason et al. (1980), Clark and Hosking (1986), Neter et al. (1993), and Foody (2004) all recommended that this test should only be used where the sample size is reasonably large. It is a general assumption in statistics that large sample sizes (n) are those greater than 30 (Eason et al. 1980; Rees 2001). As the sample sizes for both variables used here were always greater than 30 it was assumed that the use of the z-test for this investigation was valid (this assumption is corroborated by Eason et al. (1980) pg. 474 with particular reference to the z-test). Significant differences were tested for at the 5% significance level, as the test was two-tailed tab-t was given where $\alpha = 0.025$.

For assessing the success of the height modelling it was necessary to use reference datasets with which the filtering results could be compared, these included the RTK GPS points and the high resolution photogrammetrically derived DTM. Based on the assumption that the reference datasets were more accurate representations of the bare earth than the filtered laser models, then the differences between them can be used as indicators of the location and amount of model error. These differences can then be analysed quantitatively to provide information about the characteristics of the errors. The simplest method for this analysis is a qualitative investigation of the DTM, in which the surface alone is observed for phenomena such as 'dimples' or 'depressions' which will show obvious errors. Although this method is quick and simple it provides no procedure for quantitative description and it is, of course, not always possible to visually detect errors within the DTM. As an alternative, many researchers have used the differences between known and model height values at a given number of locations, and have used these to calculate the RMSE.

The use of descriptive statistics for comparing DEMs is reasonably common: examples include Fischer and Dowman (2002) and Dowman and Fischer (2003) who used the RMSE, and other descriptive statistics of the calculated differences in their investigation of errors in IfSAR (Interferometric Synthetic Aperature RaDAR) derived models. Li (1992) also advocated the use of simple descriptive statistics for modelling DEM accuracy, while Lopez (1997) based his analysis on the maxima of

the differences between the reference and testing models. The use of descriptive statistics in this way is a valid technique for assessing global error, providing that the reference data are of higher accuracy than the model DTM. However, as mentioned in Chapter 3, the RMSE does not provide any information about individual values within the DTM such as how well each cell represents the true elevation. There is a wide range of alternatives to descriptive statistics which may be used to provide more information about the spatial variation of errors within a DEM.

In many respects a spatial statistics approach would seem to be the most sensible, particularly given that errors in spatial data are spatially autocorrelated (Ehlschlaeger 1998; Wechsler 2003). For example, an error in the value of one laser point will affect its neighbours during interpolation, and will consequently affect the form of the filtered DTM. This can create systematic errors in elevation models which clearly poses a problem for non-spatial statistical methods used to define accuracy, such as the RMSE, as these methods do not accommodate the spatial autocorrelation of uncertainty. However, as was shown in Chapter 3, despite the existence of a number of techniques for spatial modelling, many of these employ methods which are too complex, or too computationally demanding, to be practically included within this investigation. As a result, the approach taken here for describing the errors within the DTMs is relatively simple - and relies on qualitative descriptions of the model differences and quantitative statistical analysis to determine whether the datasets are significantly different. In effect this method is an extension of those such as Li (1992), Fischer and Dowman (2002), and Dowman and Fischer (2003). Comparisons of the filtered datasets with the two reference datasets provide an indication of model error which is then analysed statistically, using the student t-test, to ascertain whether the results are significantly different from one another. A preliminary error surface is shown, which is simply computed from the residuals calculated during difference calculations. The methodology for the image differencing and comparison with GPS points is described below.

The filtered datasets were first compared with the photogrammetrically derived ref-

erence DTM. In order to establish differences each of the test datasets was subtracted from the reference model. The resulting images showed the areas of greatest variation between the filtered datasets and the reference model. Whilst the outputs from both the Iterative Robust Interpolation (IRI) and the geometric methods were grids and therefore straight forward to analyse as rasters, the outputs from the Adaptive TIN Method (ATM) were datasets of discrete points which could not be used for differencing in this way. In order to use a standard methodology for all datasets, and so increase the validity of the investigation, the discrete ATM points were exported as a grid at the required resolution for each output dataset. This was done using the export lattice grid option within the TerraScan software. This method creates a grid by assigning values for each cell based on the lowest elevation in each. This process may have introduced some error into the dataset, however it was anticipated that this would be minimal. In the previous chapter it was identified that the gridding process may introduce errors into datasets, however it was noted that these errors occurred near the edges of features. As the data being gridded here was a filtered, bare earth dataest it is likely that few errors were introduced. The differences were analysed qualitatively, and the results are documented in Chapter 8.

In addition to the qualititative analysis of the differences, which was based on the patterns within the images, some statistical analysis was conducted. Statistics computed included the standard deviation and RMSE which were used to provide an indication of the amount of difference between any two DTMs. Other descriptive statistics calculated included the mean, the mode, and the maximum and minimum differences. The image histograms were analysed for each difference image to provide additional information about the spread of results. The locations of Type I and Type II errors were also mapped and analysed. Type I errors are those in which bare earth points are rejected, Type II errors occur where object points are accepted as ground points.

The values in the filtered surfaces were also compared with each of the RTK GPS readings at the specified ground control points (GCPs). The location of the GCPs

Figure 4.12: Locations of ground control points at which RTK GPS readings were captured. Locations are marked by crosshairs

is shown in Figure 4.12, the sampling strategy and processing of these points was described in Chapter 4. The difference between the GPS height values and the filtered datasets was calculated, and the descriptive statistics for these residuals calculated.

4.5.5 The Effect of the Segmentation Process

The third main investigation aimed to assess the effect of the segmentation process. Segmentation was defined in Chapter 1 as the classification of neighbouring laser points, or pixels into specified 'above ground' classes. This technique is often used as a prerequisite for the 3D reconstruction for the modelling of discrete objects.

This investigation was specifically designed to analyse the effect of using different segmentation methods on the characteristics of the errors within the resultant classifications. In order to compare the effect of using different data formats, one point based and one grid based segmentation method were used. The product of each of the segmentation procedures was the classification of all above-ground objects. The quality assessment of this classification had two components. Firstly, the success of the classification itself which was assessed by manual comparison with a high resolution aerial photograph. Secondly, the success of the modelling of heights of the classified objects which was assessed by comparison with the heights captured from terrestrial laser scanning. It was considered that differences between results for each segmentation method would suggest that using different parameters would have a significant effect on the errors within resultant datasets.

As with the filtering investigation, the classification accuracies of the segmented datasets were compared by calculating contingency tables which were analysed using the z-test. Each contingency table was also used to compute producer's accuracy results to give an indication of the success of the classification for each of the individual classes. The locations of Type I and Type II errors were also mapped and analysed. Within each of the types of errors there is a distinction made indicating the type of object which has been mis-classified.

To assess how well the segmentation datasets can be used for the modelling of height, the classified points or pixels would ideally be used to compile discrete vector models of above ground features. However, this task is a complex one and has been the focus of much previous modelling research (Vosselman and Dijkman 2001; Brenner 2005; Vosselman et al. 2005), and as such is a research topic in its own right, and developing methods for assessing the accuracy of vectors and reconstruction methods is beyond the remit of this investigation. In the absence of discrete models for the 'above ground' points the only alternative was to use the heights as modelled in a normalised digital surface model (nDSM) and compare these with a reference dataset. An nDSM is a model of the height above ground (bare earth), and so

whilst it is a continuous model it does provide heights of the above ground objects. These heights may be measured in the surface model and compared with a reference dataset and the differences compared. Figure 4.13 is an example nDSM created from the nearest neighbour 1m dataset filtered with the IRI algorithm.

Figure 4.13: Example normalised Digital Surface Model

As noted above, this investigation required that the segmented points were modelled as a continuous, gridded surface. There were some restrictions on the datasets and algorithms which could be compared in this way. Clearly the raw data could not be gridded without the introduction of some interpolation errors (as is shown in Chapters 6 and 7). In addition, the investigation was only possible for the geometric algorithm (which processes grid datasets) and the IRI algorithm which outputs as a grid. Given that not all of the algorithms could be compared, only a selection of the gridded datasets were compared with the reference model to provide an indication of the general patterns of error. It is worth noting here that the irregular point datasets could have been modelled as a TIN and in this way used to create a TIN nDSM by subtracting the DTM from the DSM. However, there are problems with

TIN algebra, in the case of subtraction values to be subtracted must be defined in line with the required resolution of the derived TIN. These choices will affect the error within the resultant TIN, as will the interpolation methods used to determine the height values within the TIN at intervening locations. It was beyond the remit of this investigation to also analyse the errors within the TIN structure, but it is identified here as an area for potential further work.

The reference dataset chosen was a high resolution terrestrial laser scanning dataset. This was used to create a sample dataset of known heights on distinguishable objects. There were some constraints on the choosing of points for the sample. Firstly, the points had to be on above ground objects; secondly they also had to be on surfaces which would have been visible from the air. In terms of which objects to choose, it was decided not to use points on vegetation as there was some time delay between the capture of the two datasets, and the height of vegetation features may have changed (either just due to movement with the wind, or given the time difference between capture the vegetation may have grown). For this reason, only points over buildings were chosen. For the sample, flat roofed buildings and the peak of some pitched roofs were chosen manually as these were considered to be permanent features which were visible in both datasets. A sample size of twenty points was used, and given that there was some difficultly in finding suitable sites this was the largest practical number of sites.

The nDSMs were then compared with the reference dataset, residuals were computed, and descriptive statistics were recorded. No t-test was conducted to compare the means as the number of samples was not representative of the population. Figures 4.14 and 4.15 show the object height measurement process using the terrestrial laser data.

Figure 4.14: Showing the measurement of object heights within the reference dataset: example 1. Note the arrow shows the measurement from the base of the building (ground height) to the edge of the building roof

Figure 4.15: Showing the measurement of object heights within the reference dataset: example 2

4.6 Assessing the Adequacy of the Methodology

An applied study of this nature has not been attempted before within the laser scanning research community, and the combined methodology used in this thesis is original and as such has not been previously assessed. In order that the methodology

presented in this thesis might be used as a standard framework for quantifying laser scanner post-processing errors in any dataset, some assessment of its adequacy is also attempted within this investigation.

The adequacy of a methodology is scientifically determined by the level of its reliability and validity (Carmines and Zeller 1979; Polgar and Thomas 1991). There are two tests for the reliability (or consistency) of a methodology: these are stability and equivalence (Diamantopoulos and Schlegelmilch 1997). A methodology is said to be stable if the same results are repeatedly produced when retests under the same conditions are conducted. This is also known as test-retest reliability (Carmines and Zeller 1979). To enhance the quality of results in this investigation tests were repeated under identical conditions to ensure that similar, or in some cases the same results, were obtained. This ensured, for example, that no rounding errors were being introduced into calculations. The second component of methodology reliability is that of equivalence. This is the ability of the test to produce consistent results given different, but comparable, input data. In terms of this investigation, it was important that the results from the investigations were not specific to the study area. Partly for this reason a pilot study was conducted for a different area, and the results of the larger investigation compared and equivalence tested for.

Successful retesting does not make a methodology infallible, nor does it necessarily make it valid. Whilst *reliability* focuses on a particular property of an empirical indicator (i.e. the extent to which it provides consistent results across repeated measurements), *validity* concerns the crucial relationship between the concept and indicator (Carmines and Zeller 1979). Validity, therefore, is evidenced by the degree to which a particular indicator measures that which it is supposed to measure (Polgar and Thomas 1991). This study was designed to test the hypothesis that the choice of post-processing methodology significantly affects the amount and spatial pattern of error within ALS derived DEMs. To ensure that the methodology used in this investigation was valid a number of tests were conducted in line with Carmines and Zeller (1979) and Diamantopoulos and Schlegelmilch (1997), who all define

three different types of validity: content, criterion, and construct validity. Content validity is the extent to which a technique/methodology measures that which it is designed to measure. This can be tested subjectively, and Diamantopoulos and Schlegelmilch (1997) propose that agreement between expert judges regarding the suitability of the measure will suffice. The content validity of this investigation was tested in this manner through publication of the methodology and results in a peer reviewed journal and in peer reviewed conference proceedings, and the presentation of the methodology at external conferences to ascertain whether the measures used to quantify error in post-processing were acceptable. Comments received from external reviewers suggested and confirmed the validity of the methodology used. The second test for validity is for criterion validity, which is the extent to which a measure can be used to predict future scores of the criterion (Diamantopoulos and Schlegelmilch 1997). Criterion validity was tested for as part of this investigation by using both a pilot and a main study. In this way the pilot study was used to predict the occurrence of likely errors in the main study area. The third type of validity is construct validity, which is the examination of the results of similar tests generated using different measures. In this instance, as the errors in post processing of laser data have not been examined in this depth previously it was not possible to test for construct validity. However, where possible standard tests were employed for the quantification of error, these tests have been used previously in investigations for quantifying the success of operations. The results from the tests used here were compared to previously published results to test for equivalence.

4.7 Chapter Summary

This chapter presents the methodology used in this research, the aim of which is to both identify and analyse the errors introduced during ALS data restructuring, filtering and segmentation. A justification is offered for the choice of techniques used. The following four chapters (5-9) present the results from the investigations

described here.

Chapter 5

Quantifying the Data Capture Errors

5.1 Chapter Overview and Aims

This chapter details the first of four investigations presented in this thesis. Here, the investigation focuses on identifying the errors which are introduced during the capturing of airborne laser data. In order to differentiate between those errors introduced by post processing and those created by data capture the latter were estimated in a comparison of the laser data with a reference dataset of known (higher) accuracy. The experiment is described and the results presented below.

5.2 Quantifying the Error in Data Capture

The sources of error in data capture were identified in Chapter 2, and have previously been investigated by Behan et al. (2000) and Behan (2000) amongst others. It is not the aim of this investigation to further the work of the aforementioned authors, or to offer solutions for the reduction of error in data capture, as this is not a primary

focus of this thesis. However, the quantification of the errors in data capture is a necessary part of this investigation in order that post-processing errors can be identified and analysed.

In order to quantify the errors introduced during the capture of the airborne data a comparison was made between the airborne points and the terrestrial laser points (which were used as a control dataset). The quoted geometric accuracy of the terrestrial points in z was .25cm plus the average accuracy of the RTK GPS which was 4.5cm, giving a (claimed) total accuracy of $\pm$4.75cm. The accuracy of the airborne laser points was $\pm$15cm according to the supplier. It was thus expected that the average difference between the two datasets should lie within the range $\pm$19.75cm. Figures 5.1 and 5.2 show illustrations of subsets of the terrestrial laser data used in this investigation: in Figure 5.2 the laser points have been attributed RGB values from the associated photograph.

Figure 5.1: 3D Point Cloud for the Scan with no Colour Attributed

The comparison was made using coincident points in the two datasets meeting a horizontal separation distance of 10cm. Measurements were taken on flat asphalt planes identified in the datasets, and included road surfaces and flat roofs. These were chosen as those surfaces least likely to change between the capture of the two datasets. An overlay of the two datasets is shown in Figure 5.3.

Figure 5.2: 3D Point Cloud for the Scan. Colour for the points attributed from associated imagery

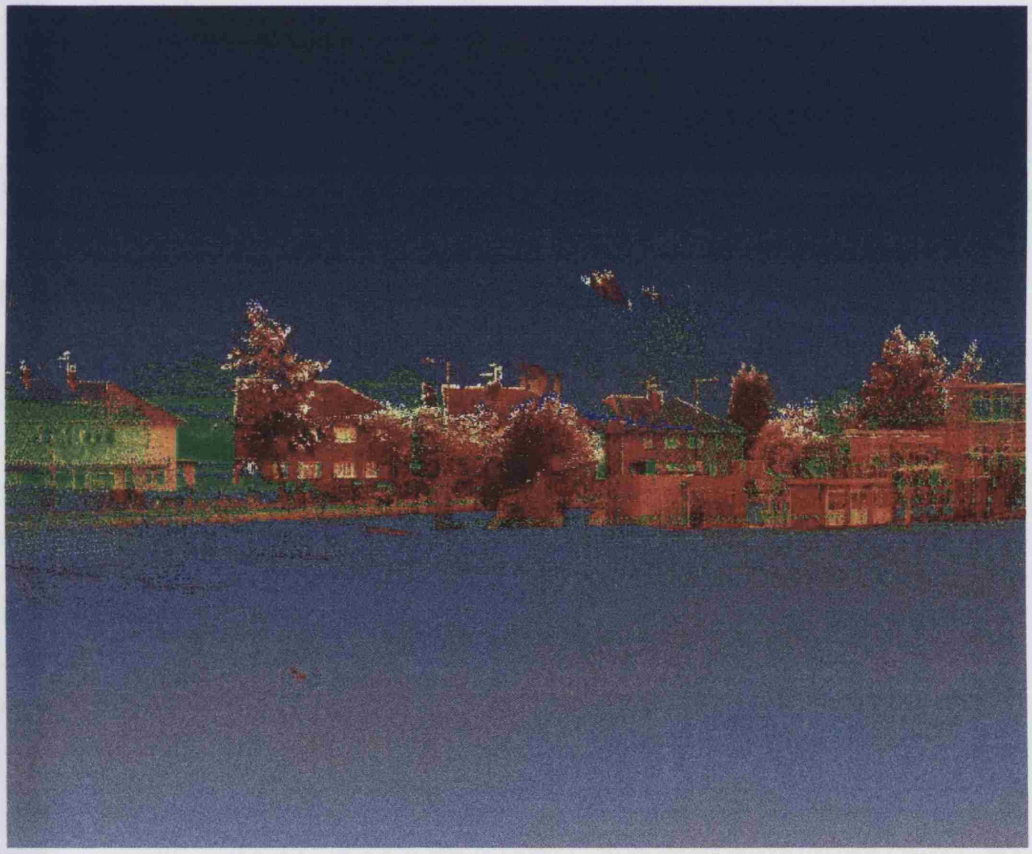

Figure 5.3: Terrestrial and airborne laser points. Terrestrial points shown in brown shades, airborne points shown in green. Note the close general correspondence between the two datasets

Table 5.1: Statistics of the Differences between the Airborne and Terrestrial Laser Data

	Type of Surface		
	Road	**Flat roof building**	**Total (m)**
Number of sample points	1002	137	1139
max	0.8390	1.2880	0.8390
min	0.0010	0.0010	0.0000
mean	0.1579	0.2188	0.1634
median	0.1320	0.2190	0.1380
mode	0.1290	0.1180	0.2040
std dev	0.1326	0.1271	0.1294

Given the requirement for coincident points to be on flat surfaces it was not possible to conduct either random or stratified sampling. The investigation required a more manual approach for picking pairs of points particularly given that dynamic objects such as cars and people had to be avoided. 1002 point comparisons were made between the two datasets for flat asphalt areas, and 137 points on flat roofs of buildings. The sample size was determined by practical resource constraints, a large number of points were sampled on flat roofed buildings, whilst it was considered that a sample size of approximately 1000 points would be sufficiently representative particularly given that samples were taken for a variety of terrestrial scans. The results are shown in Table 5.1, which shows the descriptive statistics that were calculated, in metres, from absolute elevation difference between laser points.

These measurements suggested that there was a mean difference of 16.34cm between the terrestrial and the airborne laser scanning data over both types of surface. For the purposes of this investigation, it was considered that any error calculated beyond this range would indicate the presence of error which may be attributable to post-processing. The measured errors cannot be stated categorically as being an accurate quantification of the post-processing errors because the sample of points used in the calculation of differences was not randomly selected. The sample points necessarily had to be taken from areas in which both terrestrial and airborne data points were available, these tended to be flat areas within the dataset. As such, the measured differences and errors can be used only as an indication of the presence of errors.

Based on the results calculated for this sample, the population mean (for the whole

of the Bristol study area) was estimated. The 95% confidence interval for the population mean difference, μ, is given by

$$\mu = \overline{x} \pm 1.96\frac{s}{\sqrt{n}} \tag{5.1}$$

s is the sample standard deviation (deemed to be equivalent to the unknown population standard error when divided by the square root of sample size, n), so that the range of values of the mean are:

$$\mu = (16.34 - 1.96\frac{12.94}{\sqrt{1139}}) \; to \; (16.34 + 1.96\frac{12.94}{\sqrt{1139}}) \tag{5.2}$$

$$\mu = 15.59 \; to \; 17.09 \tag{5.3}$$

So, if the distribution of errors is normal (see also Chapter 7), we can be 95% confident that μ (the average error per point for the airborne laser data) lies between 15.59 and 17.09cm. This value for predicted error in the airborne laser data capture is in line with that of the $\pm$ 15cm error quoted by the data suppliers, and corroborates previous investigations regarding the quantification of capture error such as those of Huising and Pereira (1998b) and Ahokas et al. (2003).

There are some obvious problems with this estimate; namely that the population mean error estimate is based only on the results from a non-random sample. In other words, the sample points were taken from known flat areas within the dataset, and as these areas are not representative or all surfaces within the study areas it should not be categorically stated that the population mean error truly does lie between 15.50 and 17.09cm. This said, the calculation does provide some indication of the approximate errors within the dataset.

Other limitations of the investigation relate to the difficulty of comparing datasets captured by different instruments, which are likely to have dissimilar planimetric accuracies that result in slight mismatching of the two datasets. It should be noted that the planimetric accuracy of the airborne data was not measured in this investigation as the focus of this study was upon vertical accuracy. Previous measures of planimetric accuracy include Maas (2002), and generally involve the extraction of linear features which can then be compared to their 'true' location (given by a reference dataset). The identification of objects such as buildings and vegetation forms part of this study, and further work could be conducted on the basis of these extracted features, to assess the planimetric accuracy of the datasets.

It was noted in the comparisons of the airborne data with different terrestrial scans that the calculated differences displayed a tendency to be either all positive or all negative, suggesting that some systematic error was being introduced. However the nature of this systematic error appeared not to be consistent for the whole of the airborne dataset. The discrepancy in the systematic error could potentially have been the result of differences in the corrections of flightlines or errors in transformations following excessive movement of the aircraft. Such discrepancies have been noted previously in Ahokas et al. (2003) who observed in their study that there appeared to be some degree of vertical change between flight lines in airborne data. As a result there is flight-line dependent systematic and random error affecting the total accuracy obtained.

5.3 Chapter Summary

This chapter presents a quantification of the likely errors within the raw ALS data. The calculated capture error was found to be similar to that quoted by the data supplier.

Chapter 6

Pilot Study

6.1 Chapter Introduction

This, and the following, chapter present a quantification and analysis of the errors introduced during restructuring of ALS data. This investigation was split into two distinct parts: the pilot study, and the main investigation. This chapter presents the findings from the pilot study.

The pilot study for this investigation was more exhaustive than a conventional one. It was used to develop the detail of the methodology for the main investigation, and as such employed a wider range of techniques. From the results of the pilot investigation, some of these techniques were deemed not to be suitable for urban elevation modelling and so were discounted for use in the main study. However, all of the pilot study techniques, and the results they produced, are described in this chapter in order to explicitly demonstrate why they were thought to be unsuitable.

6.2 Background to Investigation

The aim of this investigation was to quantify and map the errors introduced during ALS data restructuring. As detailed in Chapter 1, this is the process of restructuring the irregularly spaced raw ALS data into a regular data structure. Regularly spaced data points are required by a number of algorithms and software packages for conducting subsequent post-processing. In the study, the characteristics of errors introduced by different restructuring techniques were examined.

A major part of this analysis focused on assessing the effect of using different interpolation techniques for data restructuring. The comparison of various interpolation techniques has been widely researched in the past. There are some limitations of previous studies which inhibit the transferability of the results to this investigation. This is explored further below.

Interpolation is a routine process in many applications including digital image processing (Lillesand and Kieffer (2000)), cartography (MacEachren and Davidson 1987; Xie et al. 2003), soil science (Bechini et al. 2003), electronic imaging (Amidror 2002), and mathematics (Lazzaro and Montefusco 2002; Lyche et al. 2002). The effect and quality of different interpolation algorithms has been reported, based on applications in a range of disciplines (Declercq 1996; Zimmerman et al. 1999; Yang and Hodler 2000; Li and Revesz 2004; Moore 2004), including studies of how interpolation algorithms perform at different spatial resolutions (Turner et al. 1996). However these studies have not investigated the effects of interpolation onto a regular grid, nor have they examined how the interpolation effects change between different grid resolutions and very few have examined this problem in instances where the grid heights represent the urban surface.

Many studies have compared the accuracies of different interpolation algorithms. Declercq (1996), for example, investigated the prediction accuracies and portrayal of a number of different interpolation techniques. This study, like many others

(Laslett et al. 1987; Kuilenburg et al. 1982; Voltz and Webster 1990) mainly based its results on soil surface predictions or other naturally occurring and continuous phenomena. The conclusions from these studies may not be directly transferable to models of urban surface form because of the inherent complexities and irregularities of this environment. Furthermore, the lack of self-repeating structures at some scales of the urban physical environment causes a lack of predictability of surface characteristics, which leads to uncertainty for predictions of the spatial structures of errors within the surface model.

In terms of previous investigations concerning the importance of grid spacing, there has been little research into the effect of changing grid size in the interpolation stage save for that of Rees (2000) and Behan (2000). Rees (2000) investigated the interpolation of gridded DEMs. Whilst this study did not examine the interpolation of irregularly spaced data onto a regular grid, many of Rees' (2000) conclusions are nevertheless relevant, in particular that simple bilinear and bicubic interpolations are adequate for most elevation modelling requirements in non-urban areas. This conclusion is investigated in this study, in the context of an urban area. Behan (2000) also quantified error within models produced from different interpolation algorithms and found that the most accurate surfaces were created using grids which had a similar spacing to the original mean point spacing. Behan's (2000) study looked at global or average error differences between two interpolation methods. This chapter aims to extend Behan's (2000) work, by comparing additional interpolation methods, and investigating the magnitudes of error created by each method at three different grid spacings.

6.3 Pilot Study Methodology

The aim of the pilot study was to investigate a range of interpolation algorithms (using different resolutions) and to assess their suitability for modelling urban elevation in the main study. Suitability was assessed in terms of the incidence and

magnitudes of errors created by the restructuring process. The full methodology for the pilot study was described in detail in Chapter 4 and is summarised briefly below.

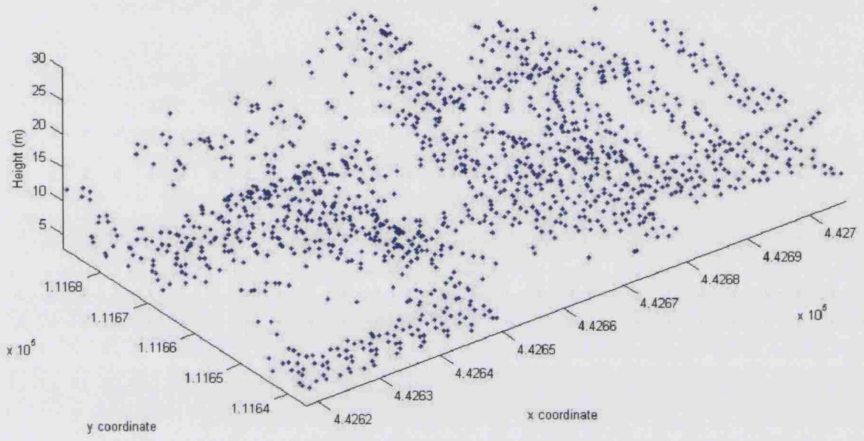

Figure 6.1: The original laser points for the pilot study displayed as a point cloud

A range of DSMs were created from a subset of a first return laser scanning dataset (Figure 6.1), supplied by the Environment Agency. The data were captured from an airborne sensor, at a point density of approximately 1 point per $4m^2$. The mean height of points within the original data was 20.02m, with a standard deviation of 15.49m. The raw points were first resampled onto a regular 1m, 2m and 4m grid, using five interpolation methods: bilinear, bicubic, biharmonic splining, nearest neighbour, and ordinary block kriging. Examples of the resultant surface forms are shown in Figures 6.2 to 6.9: some of these figures show the surfaces with the original points overlain.

There are some clear visual differences between the forms of the surfaces. The bilinear and bicubic surfaces appear reasonably similar, with the only marked difference being the incidence of some sharply defined spikes in the bicubic surface. In contrast, the nearest neighbour surface is characterised by a stepped or blocky appearance,

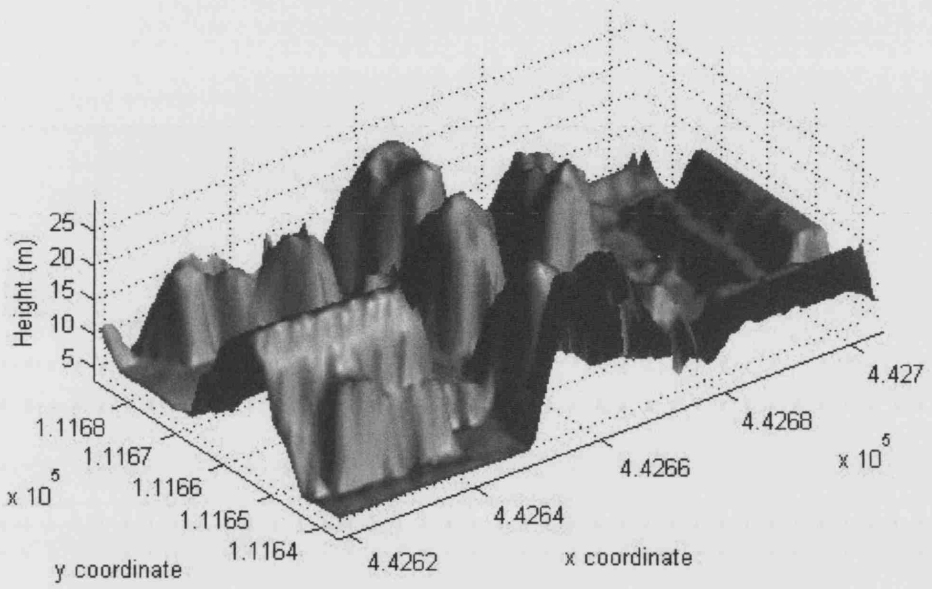

Figure 6.2: Surface interpolated using bilinear method onto a 1m grid

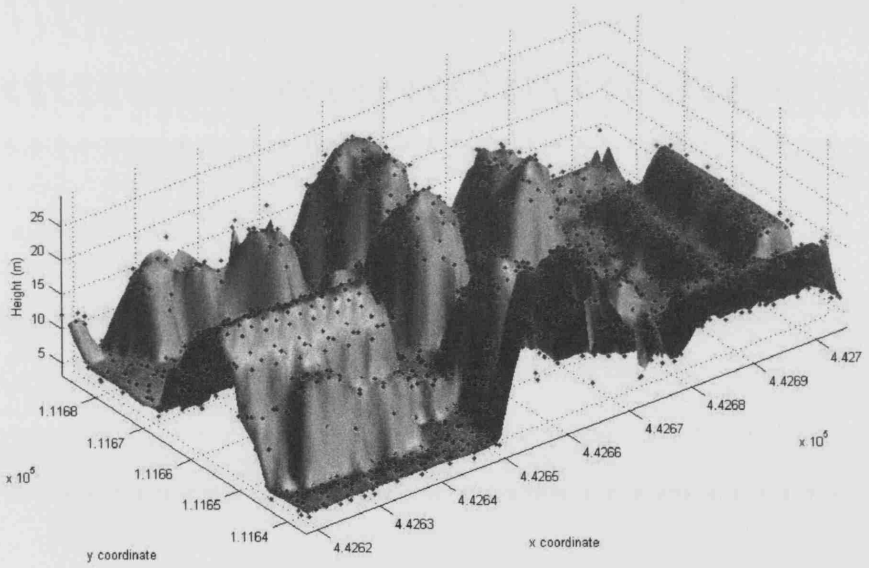

Figure 6.3: Surface interpolated using bilinear method onto a 1m grid with original laser points overlain

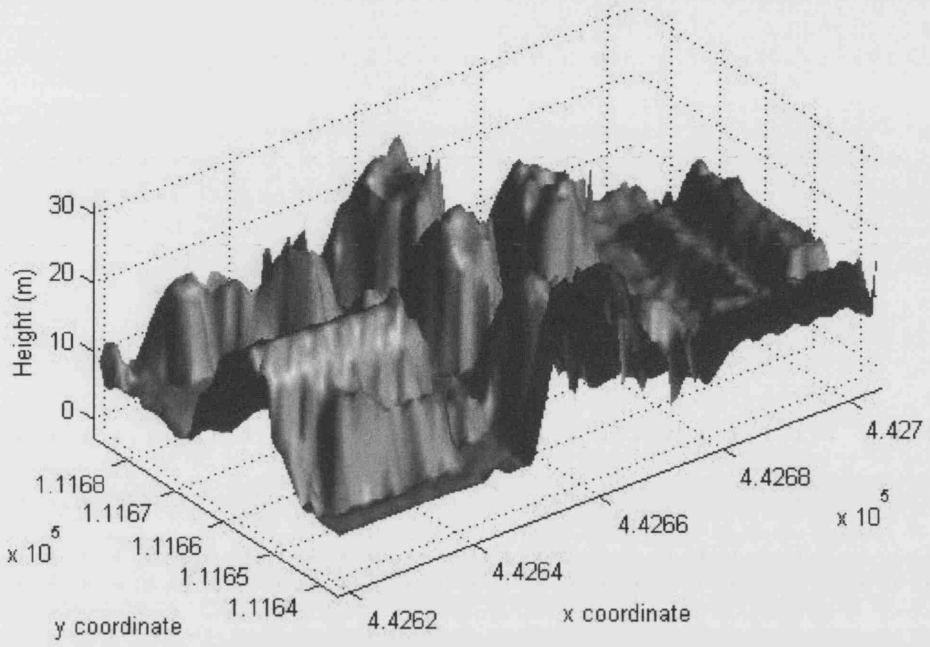

Figure 6.4: Surface interpolated using bicubic method onto a 1m grid

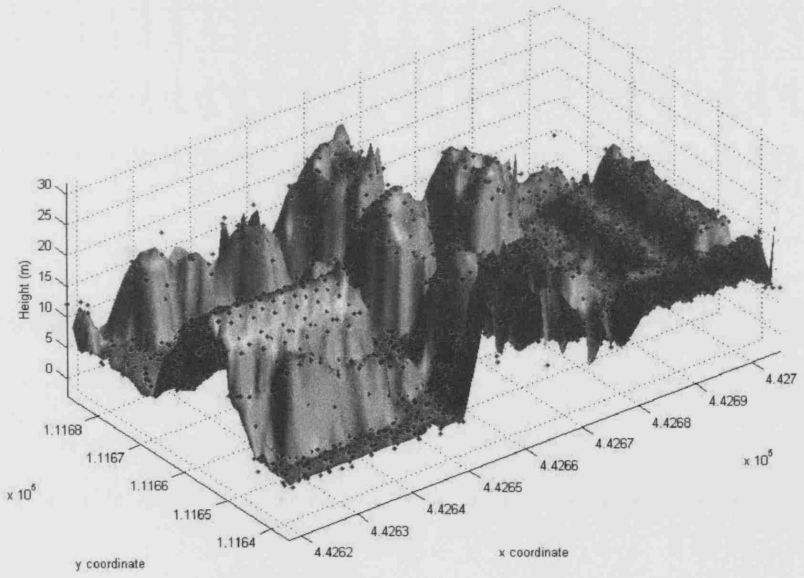

Figure 6.5: Surface interpolated using bicubic method onto a 1m grid with original laser points overlain

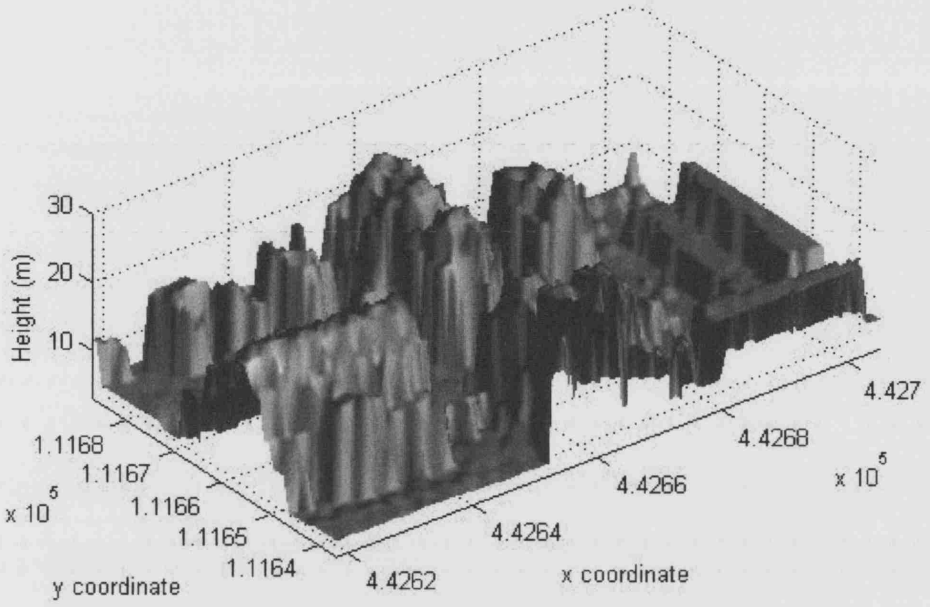

Figure 6.6: Surface interpolated using nearest neighbour method onto a 1m grid

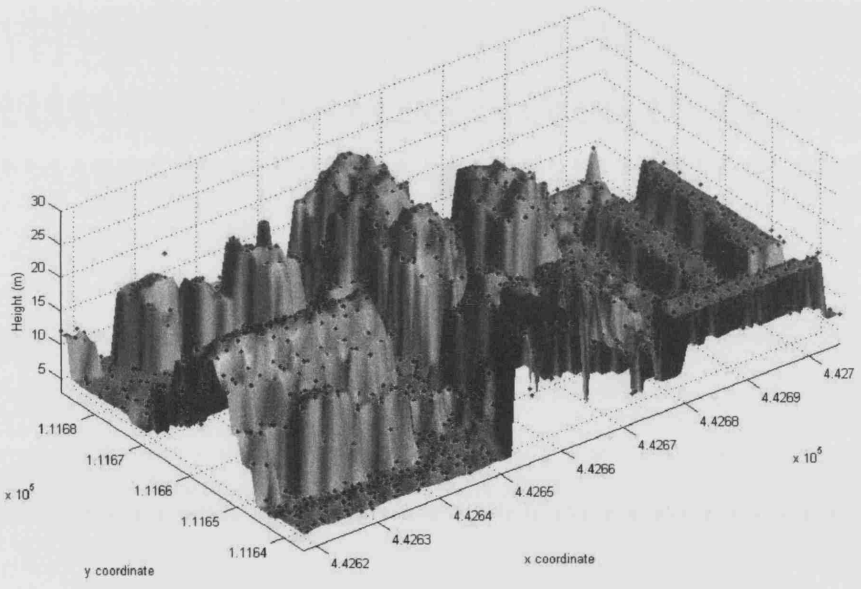

Figure 6.7: Surface interpolated using nearest neighbour method onto a 1m grid with original laser points overlain

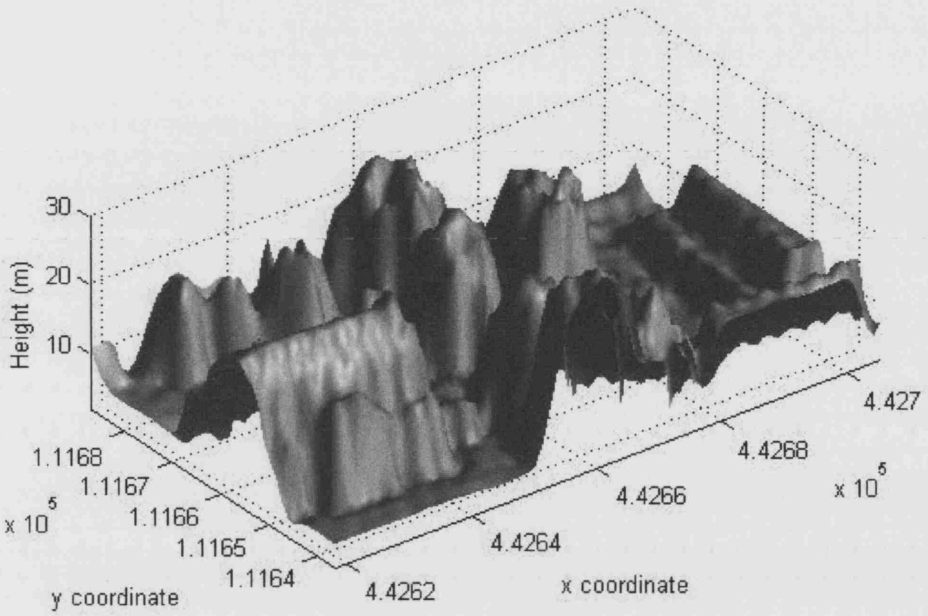

Figure 6.8: Surface interpolated using biharmonic spline method onto a 1m grid

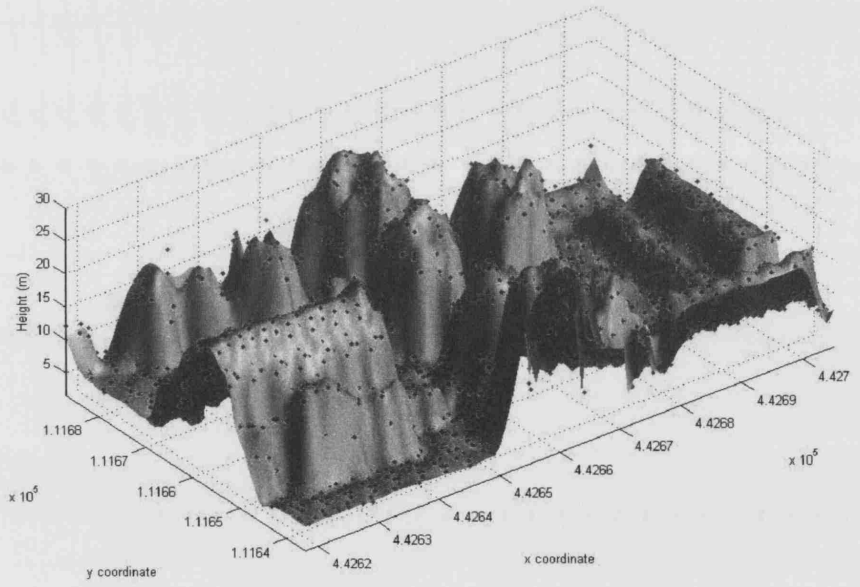

Figure 6.9: Surface interpolated using biharmonic spline method onto a 1m grid with original laser points overlain

which is a result of the algorithm's inability to model oblique surfaces. The bi-harmonic spline surface exhibits the greatest degree of smoothing, and is arguably the most visually appealing of the deterministic surfaces as it contains no abrupt changes in slope, nor does it contain any noticeable spikes.

A different surface generation method was required for kriging. In the first instance a semivariogram was computed using all of the original points to determine the interpolation weights over the whole area. A spherical model was fitted to the semivarigram, and is shown in Figure 6.10. Note that there is a pronounced dip in the semivariogram at distances greater than the range. This is termed a *hole effect* and is considered by Burrough and McDonnell (1998) to be an indication that a study region may be too small to encompass the total range of variation. This was thought to be highly likely given that this pilot investigation used such a small geographic area. The variogram range was relatively large, indicating that long-range variation dominated. The fitting of the spherical model suggests that one spatial pattern is dominant (Burrough and McDonnell 1998).

The resultant kriged surface (Figure 6.11, and Figure 6.12 with the laser points overlain) displays a very high degree of smoothing, and visually it would appear that this surface was smoother than the spline. However, a smooth surface does not necessarily equate to lower error. Indeed, qualitative evaluation of the visual appearance of the surfaces cannot provide a reliable indication of the relative success of the different algorithms. In order to accomplish this, some quantitative validation techniques were employed, these included split-sample, cross validation, and the jackknife. The results are presented in Section 6.3.1.

Each of the interpolation algorithms was also used to generate surfaces at three grid resolutions: 1m, 2m, and 4m. The reduction in spatial resolution influenced the forms of the surface, as demonstrated in Figures 6.13 to 6.15. In these figures it is clear to see that many of the features obvious in the 1m surface have been smoothed and eroded at lower resolutions. It is likely that this surface generalisation will have

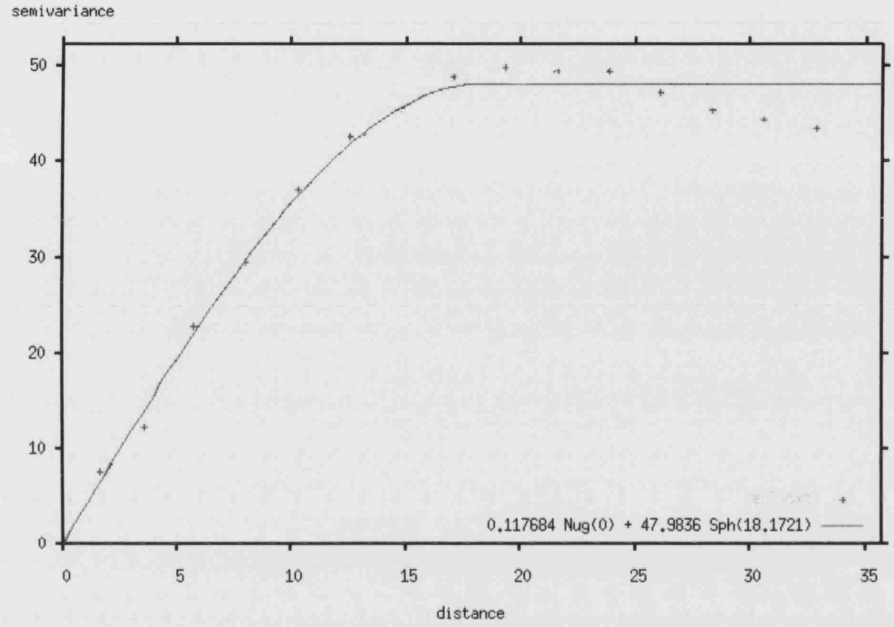

Figure 6.10: Semivarigram computed using all of the raw data point

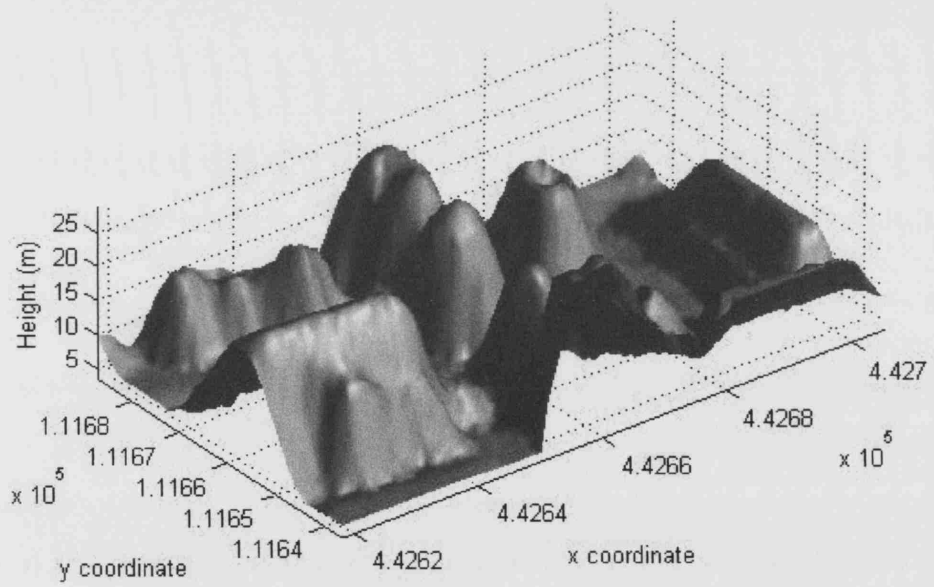

Figure 6.11: Oblique view of surface interpolated using ordinary block kriging method onto a 1m grid

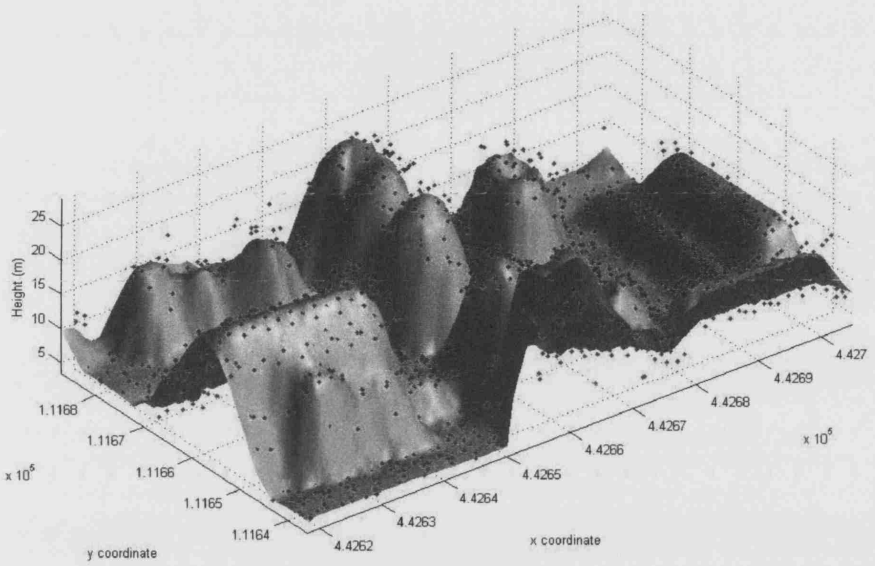

Figure 6.12: Oblique view of surface interpolated using ordinary block kriging method onto a 1m grid

introduced more error into the surface than that at higher resolutions. The extent to which the spatial resolution influenced both the amount and the location of errors in the resultant surface was analysed using the following validation techniques.

6.3.1 Results from Pilot Study Split Sample Test

The split-sample test was used to determine the stability of the methods given a diminishing number of data points. In the investigation the original laser data points were randomly split into two sets which, in the first iteration, contained 95% and 5% of the points respectively. The points in the first set (i.e. that containing 95% of the data) were used to interpolate a surface using each of the surface generation routines, at each of the three grid spacings. The values within the resultant surface were then used to predict the elevations at the locations of the points retained in the second set (5% of the points). The differences between the raw and the interpolated values were calculated. These residuals were analysed by computing the descriptive

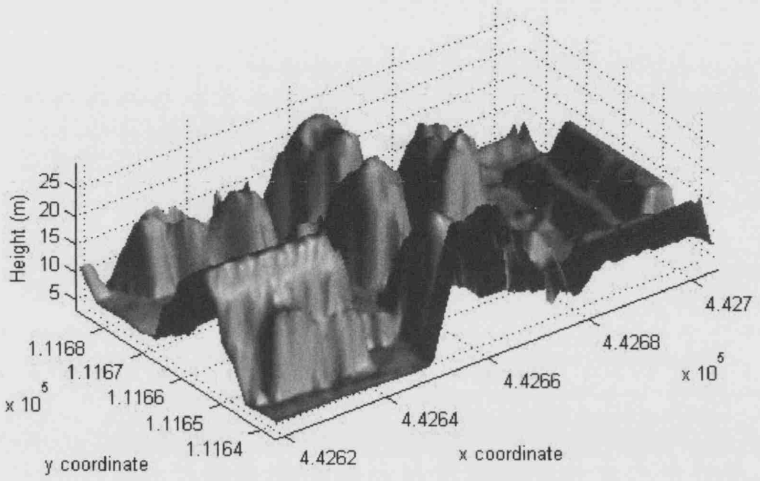

Figure 6.13: Oblique view of surface interpolated using bilinear method on a 1m grid

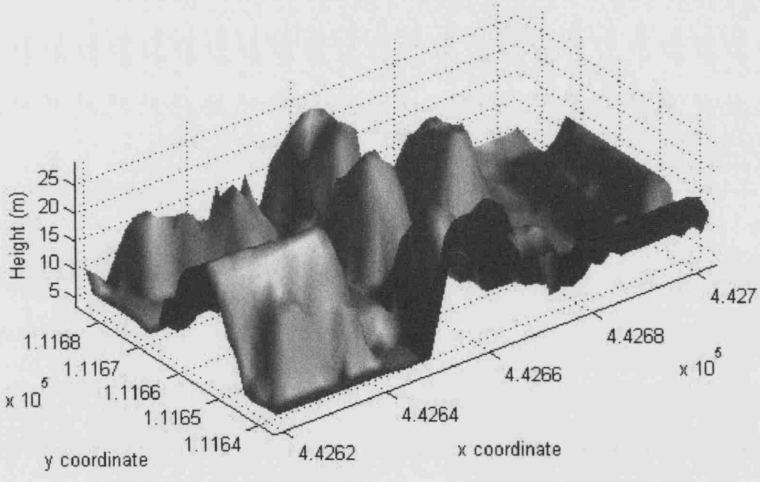

Figure 6.14: Oblique view of surface interpolated using bilinear method on a 2m grid

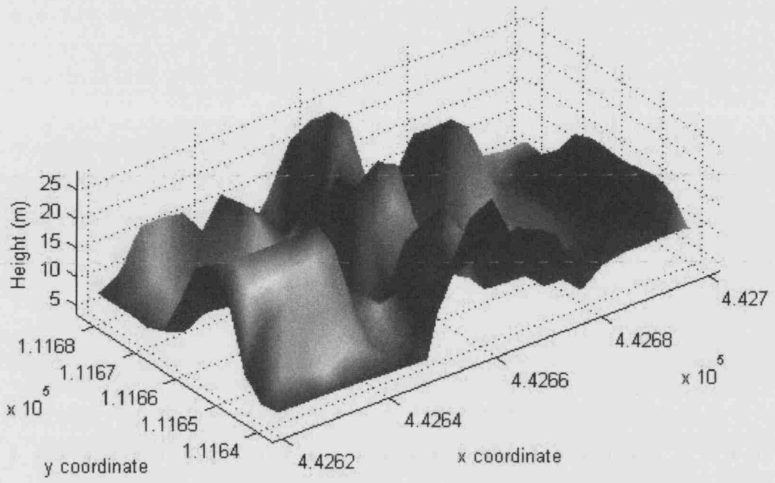

Figure 6.15: Oblique view of surface interpolated using bilinear method on a 4m grid

statistics, and by mapping the magnitudes of the errors.

In the second iteration of the split sample routine the data were split evenly, with 50% in each set, whilst in the third iteration the split was 25% and 75%. The results from each iteration were compared to assess the stability of the methods.

Table 6.1 (page 185) shows the results from a comparison of the interpolation methods for each of the three iterations. As can be seen, the bilinear and the bicubic algorithms introduced the lowest amounts of error into the surfaces in each of the three investigations with mean errors for the two methods in the region of 0.15m. The RMSEs for these, and all of the methods, were much higher than the mean error scores, being in the region of several meters. Such large restructuring errors had not been anticipated. To test the repeatability of the methodology each of the split-sample tests were run three times (using different random selections of data points for each of the two sets used in the interpolation and subsequent estimation of the error). The tests were also run using three different study areas of approximately the same size. All of the repeat experiments produced similar results and so confirmed that the restructuring process could introduce some sizeable errors into

the resultant surfaces.

Of all the algorithms compared, the nearest neighbour was found to introduce the greatest amount of error, which in some cases was twice that introduced by the bilinear and the bicubic methods. The range between maximum and minimum error for the surfaces produced by this algorithm was much larger than that for either the bilinear or bicubic methods. Similarly, the biharmonic spline and the block kriging methods produced large range values, although in terms of RMSE the amount of error introduced by either of these two methods was less than for that of the nearest neighbour algorithm.

In terms of the stability of the methods, as was expected the amount of error measured within each surface increased as the number of points for surface generation decreased from 95% to 25%. However there were some notable differences in the rate of this change, particularly for the nearest neighbour algorithm which exhibited relatively little change in RMSE values in the three experiments. Given that the amount of error in each surface was much higher than for other methods, this relative stability was not seen as a strength of this method. The bilinear, bicubic and block kriging methods all produced errors which increased by approximately 1m as the number of input points was reduced (from 95% to 25%). Errors within the spline surfaces were found to increase by only 39cm, suggesting that perhaps it was the most stable of the methods. Given fewer input points the maximum and minimum errors for the spline were found to be more pronounced, indicating that this method may produce inaccurate predictions in some regions.

As described in the previous section, the effect of altering the grid resolution during interpolation was also investigated. The grid resolutions chosen for this study were 1m, 2m and 4m. This range included spatial resolutions which were both smaller and larger than the original laser point density. The results for the split-sample test in which 95% of the points were used to generate the surfaces are shown in Table 6.2 (page 186). The results indicate that in all cases, interpolation onto

Table 6.1: Results Table for Split Sample Validation of Each Method Using 1m Resolution Data

Interpolation Method	Statistic Max error (m)	Min error(m)	Mean (m)	Std Dev (m)	RMSE (m)
95% of points used in surface generation					
Bilinear	5.39	-6.86	-0.14	1.92	1.90
Bicubic	5.20	-3.09	-0.10	1.69	1.65
Nearest Neighbour	13.34	-11.63	0.14	3.51	3.48
Biharmonic Spline	7.18	-12.78	-0.32	2.67	2.67
Ordinary Block Kriging	9.73	-6.97	0.16	2.48	2.47
50% of points used in surface generation					
Bilinear	10.18	-14.90	0.13	2.94	2.94
Bicubic	12.36	-15.71	0.17	3.05	3.05
Nearest Neighbour	15.37	-15.40	0.24	3.46	3.47
Biharmonic Spline	15.17	-15.74	0.22	3.06	3.06
Ordinary Block Kriging	14.27	-8.81	-0.06	3.08	3.08
25% of points used in surface generation					
Bilinear	13.39	-15.18	0.16	3.00	3.00
Bicubic	11.38	-16.72	0.10	2.99	2.99
Nearest Neighbour	15.99	-16.14	0.16	3.66	3.66
Biharmonic Spline	11.77	-18.24	0.19	3.05	3.06
Ordinary Block Kriging	12.00	-16.14	-0.34	3.69	3.70

Table 6.2: Results Table for Split Sample Validation of Each Method at Different Grid Spacings

Interpolation Method	Grid Spacing	Statistic Max error (m)	Min error(m)	Mean (m)	Std Dev (m)	RMSE (m)
Bilinear	1m	5.56	-7.14	-0.02	1.98	1.96
	2m	5.24	-7.26	-0.16	1.93	1.92
	4m	4.29	-7.44	-0.73	2.12	2.22
Bicubic	1m	6.13	-6.30	0.02	2.64	2.01
	2m	5.83	-6.57	-0.15	1.96	1.94
	4m	2.94	-7.50	-0.49	2.02	2.05
Nearest Neighbour	1m	8.40	-7.14	0.29	2.58	2.58
	2m	8.40	-6.94	0.36	2.41	2.42
	4m	8.33	-14.72	-0.06	2.67	2.65
Biharmonic Spline	1m	7.30	-6.27	0.13	2.18	2.17
	2m	9.28	-6.15	0.16	2.40	2.39
	4m	13.09	-22.73	-0.28	3.93	3.91
Ordinary Block Kriging	1m	6.65	-6.59	0.01	1.88	1.87
	2m	6.52	-6.28	0.00	1.84	1.82
	4m	7.55	-5.12	0.10	1.97	1.96

the 4m grid resulted in higher errors, because of the loss of information in this approach. The bilinear, bicubic, and biharmonic splining interpolators produced relatively stable range and mean errors, and there appeared to be only a minimal difference between the surfaces created at different spatial resolutions. The nearest neighbour interpolator produced higher errors at larger grid spacings than any of the other methods. Overall, the resolution which generally produced the lowest error was the 2m spacing which was the closest to the original mean point spacing. This corroborated the findings of Behan (2000).

The location of errors was briefly investigated for the pilot study. The pattern of occurrences of error in the bicubic and biharmonic spline surfaces are shown in Figures 6.16 and 6.17. In the figures the errors are viewed from above, the location of errors is indicated by a disc and the magnitude of error by the size of the disc. In general terms the patterns of errors for each of the interpolation methods, at each of the spatial resolutions, was very similar. To a large extent the incidence of the

largest errors appeared to mirror the occurrence of surface discontinuities such as the edges of buildings or vegetation as identified from the aerial photograph of the study region (see page 119). Of particular note within these errors was the clear boundary of the church, and the apparent clusters of higher errors in the region of the vegetation.

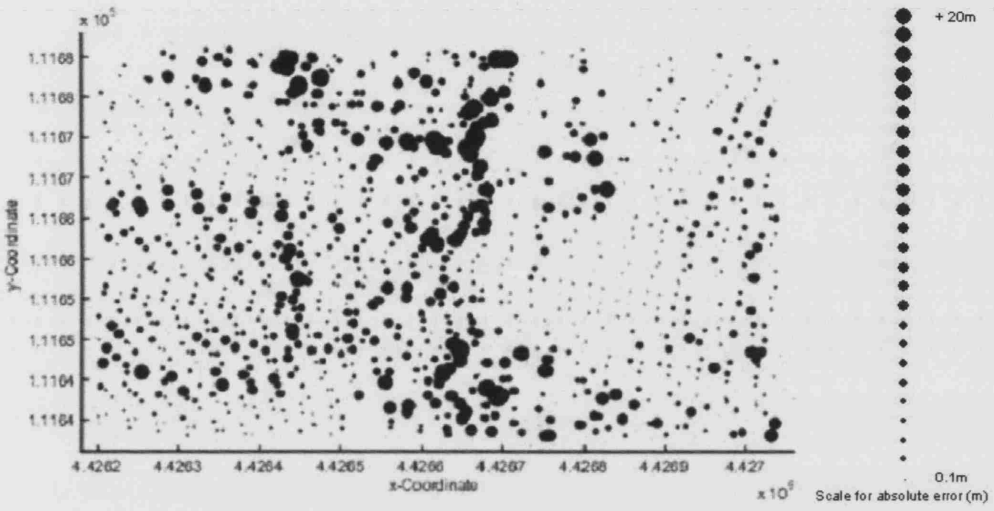

Figure 6.16: Spatial Pattern of Errors on the Bicubic 1m Surface

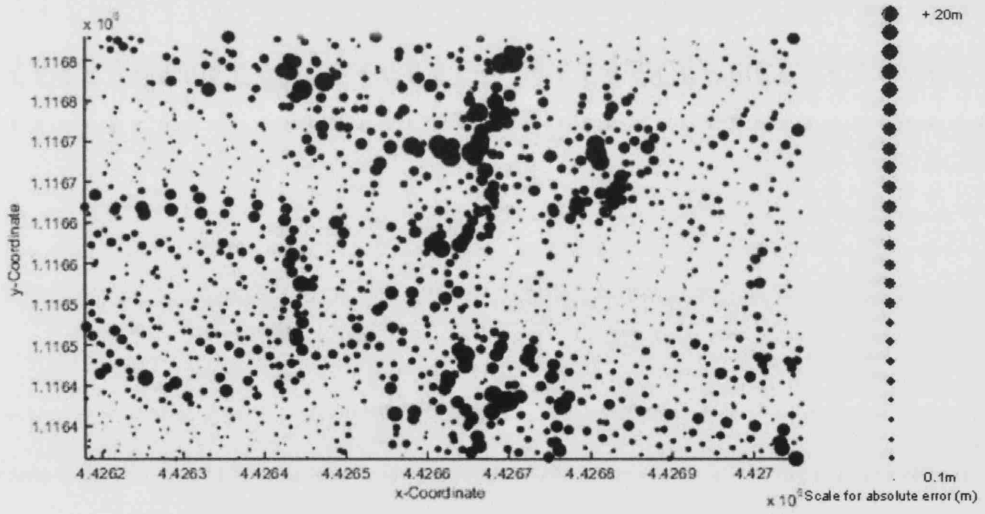

Figure 6.17: Spatial Pattern of Errors on the Biharmonic Spline 1m Surface

There were some small differences in the spatial patterns of errors between the dif-

ferent interpolation methods, namely that in the spline surface some of the largest errors appeared to occur at the edges of the dataset, indicating that there may be some spatial dependence in the occurrence of the highest magnitude errors. Problems with splines in unconstrained regions (such as at the edges of a dataset) were noted in Chapter 4. Given that it seemed likely that some of the largest errors in the spline surfaces were caused by edge effects the experiment was repeated using a 4m buffer zone around the edge of the dataset. In the modified experiment errors within the buffer zone were not included within the statistical analysis. The descriptive statistics for the errors in one of the spline surfaces both with and without a buffer zone are presented in Table 6.3. The results show that there is a general improvement in overall error for the spline surface when using the buffer zone, and that for the lower resolution surfaces the improvement in error was striking, with a reduction in RMSE from 3.92m to 2.53m. Surfaces using the other interpolation algorithms were also generated using buffer zones, although there was found to be no difference in the amount of errors generated, suggesting that it was only the biharmonic algorithm which encountered edge effect problems.

The improvements in accuracy offered by the use of a buffer zone with the biharmonic splining for this experiment suggest that it may be a more useful technique for surface generation than first indicated. However, using a buffer zone during interpolation increases processing time and may therefore be costly to the user. In addition, some manual intervention may be required for some datasets in order to determine an appropriate size for the buffer. These conditions may make the spline approach less attractive. Given that even by using a buffer zone, there was no improvement over the bilinear and bicubic techniques offered by the modified spline approach, the biharmonic spline algorithm was not considered to be suitable for use within the main investigation.

In addition to using the split-sample method to analyse the error, two further techniques were used: cross validation and jack-knifing. The results from these tests were compared with those presented above to test for consistency.

Table 6.3: Descriptive Statistics for the Spline Surface With and Without a Buffer

Biharmonic Spline		Max Error (m)	Min Error (m)	Mean Error (m)	Std Dev (m)	RMSE (m)
With buffer	1m	6.00	-5.92	0.10	2.14	2.12
	2m	5.99	-5.56	0.12	2.03	2.05
	4m	5.31	-8.27	0.05	2.55	2.53
Without buffer	1m	6.51	-11.61	-0.21	2.57	2.56
	2m	6.36	-7.53	-0.15	2.33	2.32
	4m	7.32	-23.51	-0.59	3.91	3.92

6.3.2 Results from Cross Validation

The methodology for cross validation was presented in Chapter 4. To recap, in the cross validation process one point is omitted from the dataset, the surface generated and the values used to predict the height of the omitted point. The whole process is repeated until every sample point has been removed in turn. This procedure was performed for each of the deterministic interpolation methods at the three grid resolutions, and the results are presented in Table 6.4. This validation technique was not necessary for the kriged surface as this technique provides an estimate of uncertainty for each interpolated value (discussed further in Section 6.3.4).

Table 6.4: Pilot Study Results of Cross Validation Analysis

Interpolation Method	**Grid Size**	Max error (m)	Min error (m)	Mean (m)	Std Dev (m)	RMSE (m)
Bilinear	1m	15.82	-10.85	-0.07	2.43	2.34
	2m	15.38	-10.38	-0.08	2.48	2.53
	4m	15.08	-9.84	0.09	2.75	2.65
Bicubic	1m	16.06	-13.00	-0.10	2.56	3.16
	2m	15.90	-11.65	-0.07	2.66	2.12
	4m	15.41	-7.86	-0.03	2.64	2.63
Nearest Neighbour	1m	15.56	-15.99	-0.06	3.53	1.99
	2m	15.56	-15.32	0.19	3.59	6.87
	4m	16.00	-15.19	0.20	4.14	7.27
Biharmonic Spline	1m	16.21	-13.67	-0.05	2.61	1.74
	2m	16.16	-13.13	-0.05	2.56	1.65
	4m	16.89	-13.81	0.11	2.62	3.99

The cross validation results largely mirrored those from the split-sample study, with

the bilinear and bicubic methods again producing similar amounts of error. The biharmonic spline surface at the lowest grid resolution (4m) was again found to contain much higher overall error. It was observed that the error in the nearest neighbour surface was substantially more than that in the other surfaces, however the magnitude of this error was much higher than had been calculated using the split-sample technique. Finally, in terms of the range of maximum and minimum errors, the cross validation results for all of the methods exhibited similar results.

In order to show additional characteristics of the surface errors, the interpolated values ($Z^*(i)$) for each the methods were plotted against the raw values for the three grid spacings. The resultant scatter plots for the bilinear and the nearest neighbour surfaces at 1m are shown in Figures 6.18 and 6.19. The spread of the plots and the linear correlation coefficients confirm that there is relatively good overall agreement between the predicted and the measured values for all methods. The nearest neighbour method has a lower correlation coefficient for all spatial resolutions than any of the other methods (0.804 - 0.855) which all produce values in excess of 0.9. The spread of points for the nearest neighbour plots also shows large differences between the predicted and the observed values for heights below 20m. However, it indicates that there are also many points which are predicted with a high accuracy for heights below 10m.

As with the split-sample technique, the errors were again plotted and the spatial pattern qualitatively assessed. Similar patterns were noted for all of the methods, and the apparent coincidence with breaklines was again noted.

6.3.3 Results from Jackknifing

The jackknife procedure (described in full on page 149) was used to further assess the error by analysing the residuals in the estimation of a chosen statistic for each of the deterministic interpolation techniques. The differences in the bias and error of each of the methods was used as an indication of the characteristics of the interpolators.

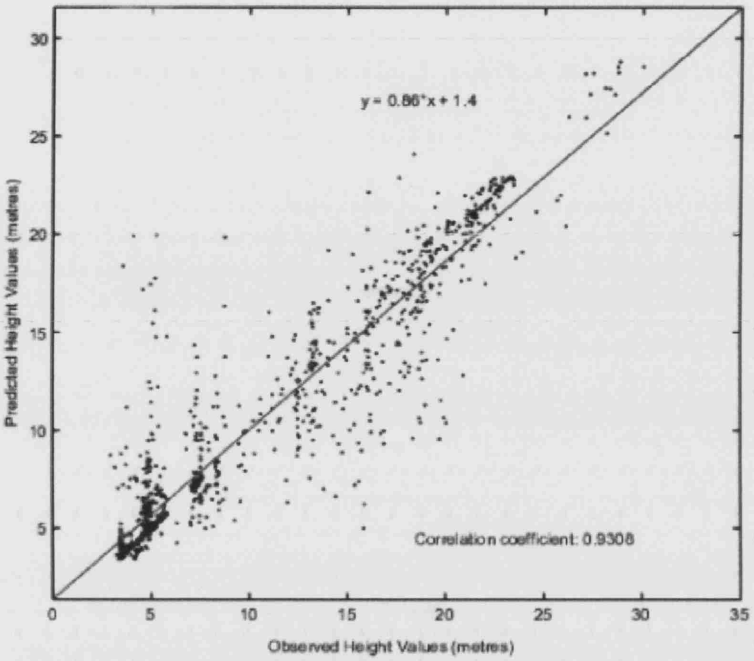

Figure 6.18: Observed vs. predicted height values using the bilinear interpolation method at 1m grid resolution for the 'leave one out' cross validation method

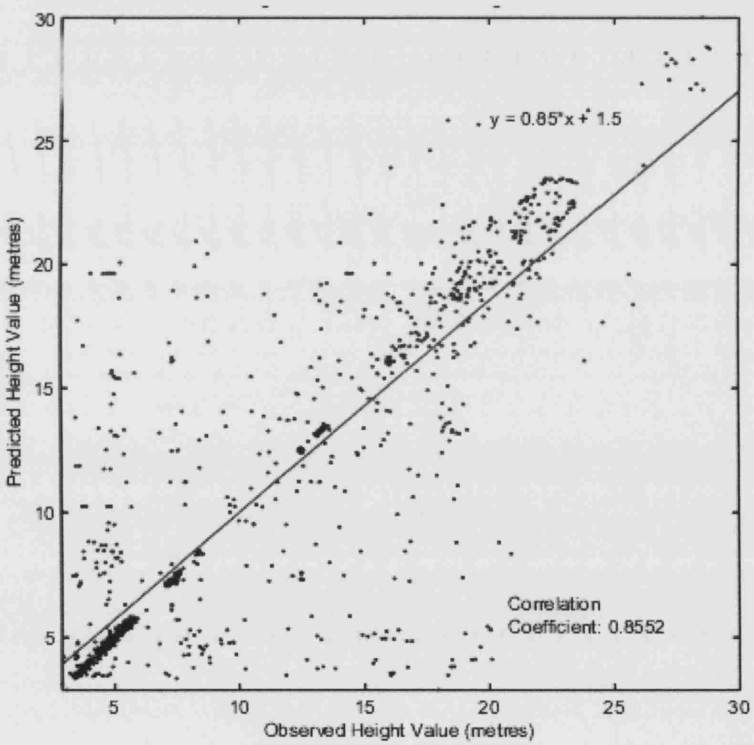

Figure 6.19: Observed vs. predicted height values using the nearest neighbour interpolation method at 1m grid resolution for the 'leave one out' cross validation method

The statistic of interest which was tested for standard error ($\hat{SE}$) was the mean ($\overline{x}$) of the interpolated surface heights. It was considered that smaller variations in surface mean would indicate a more stable surface which was less influenced by individual observations than a surface which produced a higher bias and standard deviation for the mean. The results are shown in Table 6.5.

Table 6.5: Pilot Study Results for Jackknife Technique

	Jackknife Standard Error		
Interpolation Method	**Grid Resolution**		
	1m	2m	4m
Bilinear	0.057	0.0733	0.0659
Bicubic	0.0444	0.0549	0.0443
Nearest Neighbour	0.089	0.0994	0.0103
Biharmonic Spline	0.0003	0.0004	0.0007

The results indicate that the biharmonic spline method produced the smallest overall standard error, whilst the other methods generally produced higher standard errors. It was observed that the bilinear and nearest neighbour methods produced slightly higher $\hat{SE}$ values than the bicubic method. This was thought to have been caused by the fact that the bilinear and nearest neighbour techniques create more pronounced edges in the surface which will have influenced the calculation of the mean.

6.3.4 Results from Kriging Variance

The cross validation and jackknife procedures were not used for analysing the errors within the block kriging surfaces, because kriging produces both interpolated values and an estimate of uncertainty of this value given by a variance value (σ_e^2) for each point. This can reveal important information about the reliability of the interpolated value over the area of interest. Often the kriging variance is mapped as the kriging standard error as this has the same units as the predictions (Burrough and McDonnell 1998), and this convention is followed here. The resultant uncertainty

map for the 1m surface is shown below in Figure 6.20 which displays a plaid, or chequered, pattern in the kriging standard error and no obvious evidence of any other spatial pattern of uncertainty. This is because kriging variance is related to data configuration - in areas where there are a lot of data the kriging variance is low, where there are few data the variance is high. The chequered effect thus reflects the ALS capture pattern, and the lines of points as the scanner oscillates across each sweep. The variance maps were not thought to reflect any patterns in addition to the systematic one identified above.

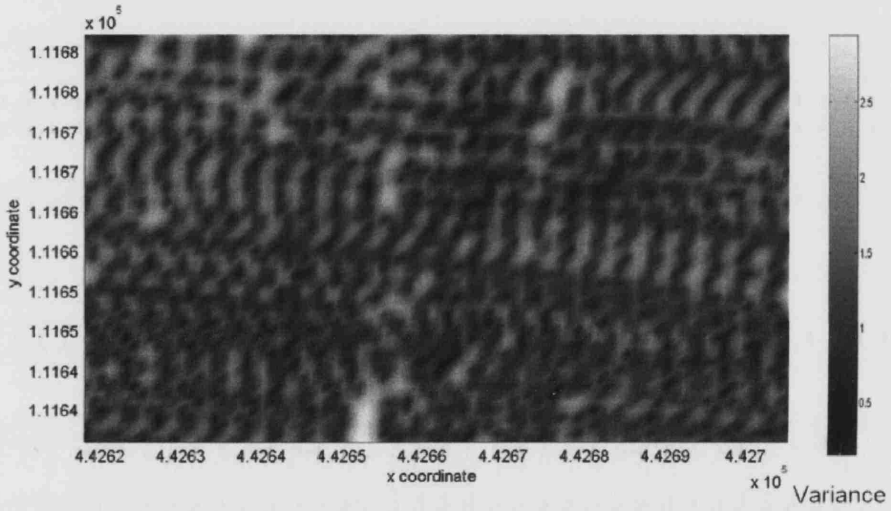

Figure 6.20: Map of Kriging Standard Error for 1m Interpolation Grid.

6.3.5 Summary of Pilot Study Findings

This chapter has used a range of approaches for three specific, but interrelated, objectives: (i) to determine which interpolation technique and which grid resolution produced the least error; (ii) which of the methods were suitable for use in the main investigation; (iii) to assess the patterns of the interpolation errors. The results obtained from each objective are summarised in turn.

In terms of comparison of the interpolation techniques, the nearest neighbour algorithm was found to produce the greatest overall error. This probably arose because of its inability to model the oblique surfaces that are common in urban roofscapes. In the nearest neighbour surface, changes in height are stepped, creating discontinuities which are not necessarily present in the raw data. Whilst introducing error, the stepped appearance also means that this method does not produce the most realistic looking representation of the urban surface. This may be an important consideration for the data user. In contrast, the lowest errors were created by the bilinear and bicubic methods with relatively little difference between the results for these two methods.

The biharmonic spline produced the largest magnitude errors, this was thought to be caused by the production of strong artificial oscillations in unconstrained regions. In other words, the spline produced large overshoots in areas of sparse data such as at the edges of the dataset and where point densities were lower. Whilst the spline produced a generally smooth surface inside the buffer areas identified, it was thought not to be a suitable technique for use in the main investigation due to practical limitations - such as the time taken to define the buffer zone for each dataset.

The ordinary block kriging surface was found to introduce more error than either the bilinear or the bicubic method. This may indicate that kriging should not be used for modelling artificial structures such as those in the urban surface. Kriging has traditionally been used for modelling naturally occurring phenomena, such as mineral outcrops for the mining industry. The transferability of kriging for modelling the highly discontinuous, anthropogenic urban environment requires further investigation. It should be noted that the results presented here relate to the specific modelling choices made, such as the variogram, the search radius, and the maximum/minimum data for kriging, rather than the technique itself. Further studies using different parameters, and different types of kriging would help determine the suitability of this as a general technique for modelling urban elevation surfaces.

These additional investigations were not conducted within this research following the advice of Dr Dan Cornford (Personal Communication, September 2003) who indicated that differences between the surfaces produced by different kriging techniques tend to be very subtle. In view of the limitations of block kriging identified in this investigation, geostatistical interpolation was not deemed to be suitable for the main investigation.

The incidence of high magnitude errors appeared, qualitatively, to mirror the occurrence of surface discontinuities. Similar observations have previously been noted in the modelling of other surfaces, and a number of strategies have been proposed for reducing the errors in these locations (Vable (2001) and Florinsky (2002) both advocated an understanding of the Gibbs phenomenon, which is the specific behaviour of some functions near discontinuities). The pattern of errors over discontinuities, such as those identified here, could be used as an input for subsequent modelling stages, such as feature extraction, as it was noted that there was a difference in the amounts of errors over different breaklines, with the errors caused by building breaklines tending to be smaller than those caused by vegetation. The pattern of errors could also be used in the reduction of error in the interpolation process by using a two-stage process. The spatial pattern of errors could be calculated and used to identify breaklines in the first stage of modelling. This breakline information could then constrain the Delaunay Triangulation in the second stage of interpolation, thus reducing the errors.

6.4 Chapter Summary

This chapter has presented the results from the pilot study for the investigation regarding the characteristics of the errors introduced by data restructuring. The following chapter builds on these findings and applies the adapted methodology to a much larger study area.

Chapter 7

Analysing the Errors Introduced during Data Restructuring

7.1 Chapter Introduction

The results from the pilot study (see Chapter 6) were used to define the most appropriate methodology for the data restructuring investigation over a larger study area. This chapter describes the revised methodology and presents the findings from the main study area. The chapter begins with a brief resume of the nature and extent of the main study area site.

7.2 Main Investigation

7.2.1 Study Site

A full description of the study area was presented in Chapter 4. To recap, the study site was a 1km^2 block near Bristol, SW England. This area (Figure 7.1) contained a variety of land types, including urban residential areas, bare fields, and sloping

vegetated terrain. This diversity helped to ensure that the results were not specific to any particular type of terrain. The site was relatively flat, with a rise in gradient towards the south-west. The mean height of laser points for this study region was 20.02m, with a standard deviation of 15.49 m.

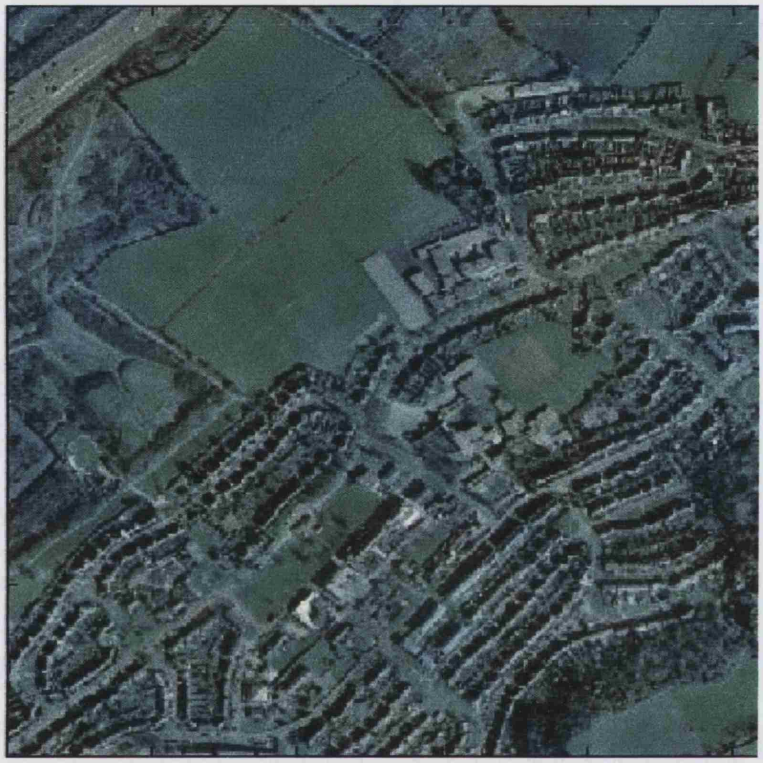

Figure 7.1: Orthorectified aerial photograph of the 1km^2 Bristol study site

The findings of the pilot study were used to devise a strategy for ALS data restructuring on a much larger scale for the main study area. Given the limitations of both the biharmonic spline and the block kriging methods these algorithms were eliminated from the main investigation, and the analysis focused on the bilinear, bicubic and nearest neighbour algorithms. In terms of grid resolution a similar approach to that of the pilot study was used, in which grid sizes both smaller and larger than the original point spacing were chosen. The point spacing of the ALS data for the main investigation was approximately 1.2m, and the associated grid resolutions chosen for the investigation were 0.5m, 1m, and 2m.

To quantify the errors within the resultant surfaces only the split-sample routine was

adopted. This was largely because of practical limitations related to the extensive processing time required for the cross validation and jackknifing techniques of the main study area (see Chapter 4, page 151, for further discussion). It was noted in the pilot study, that using only the split-sample routine alone can be problematic as the results it produces are dependent on the values randomly chosen for surface generation. Some attempt was made to investigate the extent of this problem by running repeat experiments using different randomly selected values. It was found that reasonably similar error values were produced each time, suggesting that even using the split-sample method alone could provide some indication of the errors incorporated during restructuring.

Some examples of the resultant surfaces for the main study area are shown in Figures 7.2 to 7.5. These may be compared with the orthorectified aerial photograph (Figure 7.1) for context. All of the resultant surfaces are presented in Appendix D. There is little obvious difference in the surface forms viewed in this way, partly because of the size of the study area and the small scale of the changes between the surfaces. To analyse and quantify the differences between the errors within the surfaces, some further analysis was required.

In order to quantify the errors within the generated surfaces the split-sample techniques, again using 95%, 50%, and 25% of the data, were used. The results for each of the three interpolation methods at each grid resolution for each of the three split-sample tests are presented in Tables 7.1 to 7.3 (page 201). Findings from these results can be split into those relating to the overall performance of the methods, the stability of the methods, and the effect of changing grid resolution on the nature of the resultant error.

Firstly, in terms of the overall performance of the algorithms, it was clear that the bilinear surface produced the smallest error, thus corroborating the findings from the pilot study. However, the descriptive statistics for the bicubic method differed significantly from the preliminary investigations. The huge range of values between

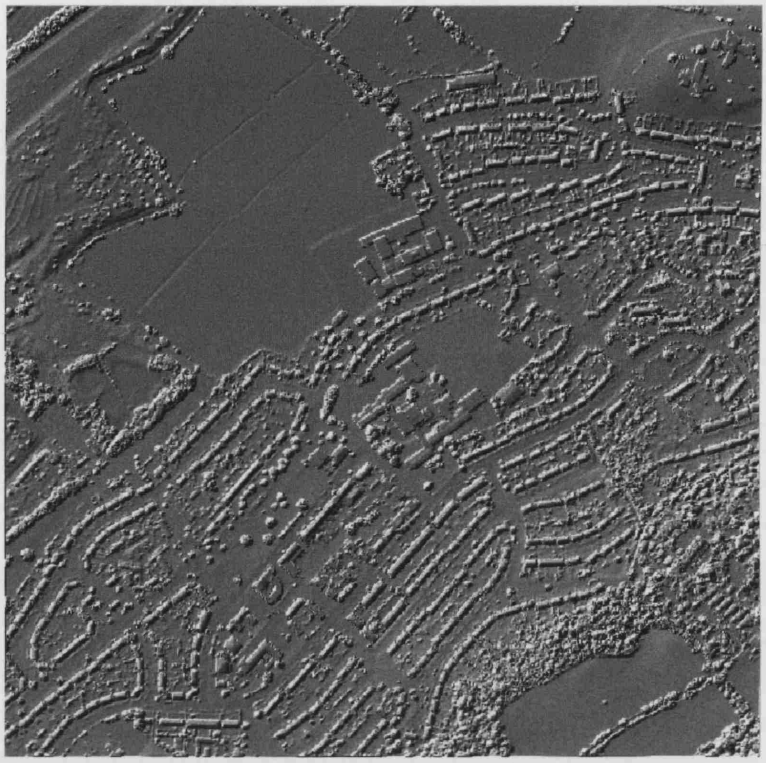

Figure 7.2: View of the hill shaded surface created using the bilinear algorithm at 1m resolution

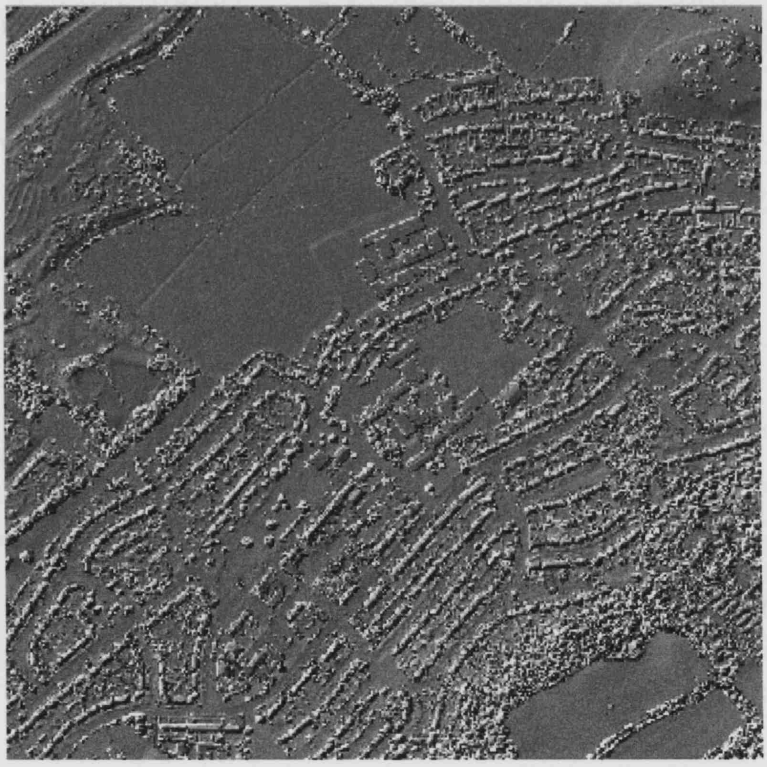

Figure 7.3: View of the hill shaded surface created using the bicubic algorithm at 1m resolution

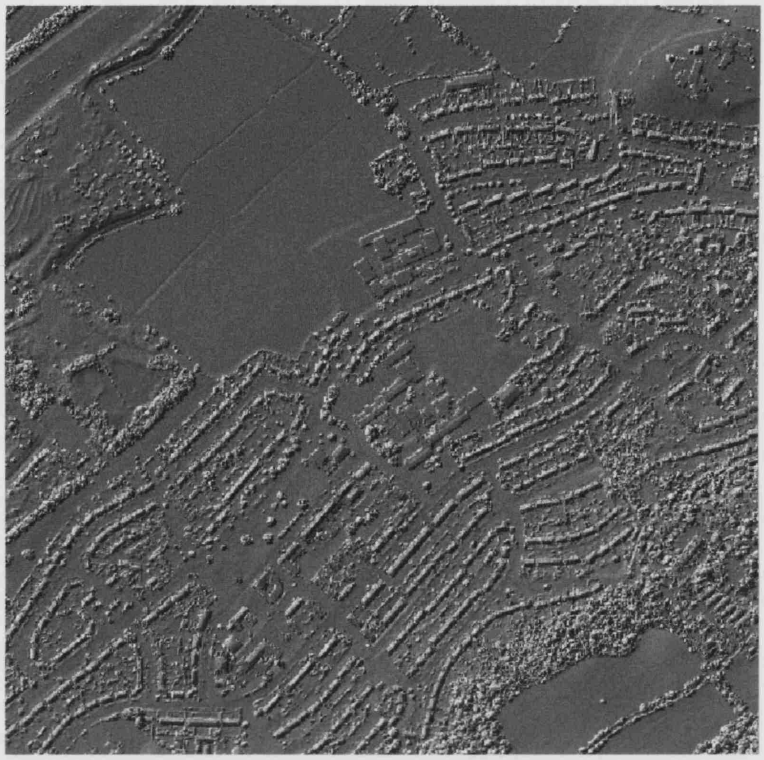

Figure 7.4: View of the hill shaded surface created using the nearest neighbour algorithm at 1m resolution

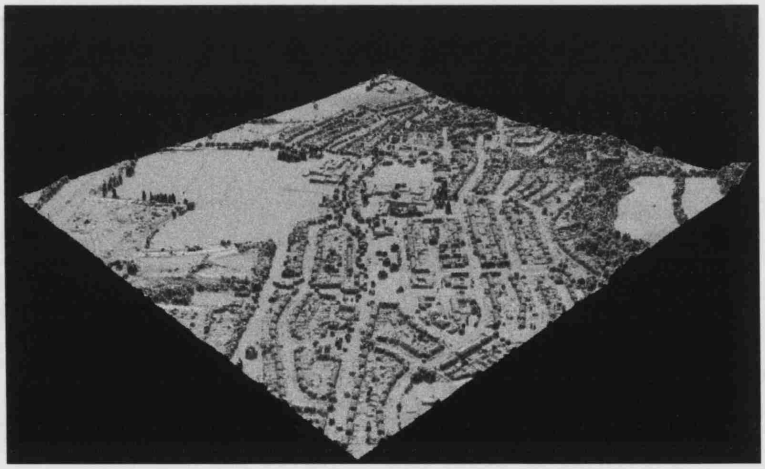

Figure 7.5: Surface profile of the nearest neighbour 1m surface

Table 7.1: Results showing the errors identified using the split-sample test for the main investigation. Surfaces generated using 95% of the data

Algorithm	Resolution	Statistic max	min	mean	std dev	RMSE
Bilinear	0.5m	18.31	-22.74	0.00	1.29	1.29
	1m	17.28	-22.83	0.00	1.25	1.25
	2m	15.63	-22.69	0.01	1.28	1.28
Bicubic	0.5m	3201.00	-505.14	0.04	13.69	13.69
	1m	392.43	-536.29	0.00	4.93	4.93
	2m	464.57	-202.53	0.03	3.91	3.91
NN	0.5m	25.32	-23.42	-0.01	1.64	1.64
	1m	25.32	-24.15	0.00	1.65	1.65
	2m	22.91	-23.42	0.01	1.70	1.70

Table 7.2: Results showing the errors identified using the split-sample test for the main investigation. Surfaces generated using 50% of the data

Algorithm	Resolution	Statistic max	min	mean	std dev	RMSE
Bilinear	0.5m	21.31	-25.21	0.00	1.31	1.31
	1m	20.00	-25.14	0.00	1.28	1.28
	2m	20.73	-24.74	0.00	1.29	1.29
Bicubic	0.5m	8925.90	-2993.50	0.01	14.24	14.24
	1m	931.72	-431.02	-0.01	3.66	3.66
	2m	269.94	-303.74	-0.01	2.25	2.25
NN	0.5m	24.95	-25.42	-0.01	1.64	1.64
	1m	25.68	-25.42	-0.01	1.65	1.65
	2m	24.36	-26.30	0.00	1.73	1.73

Table 7.3: Results showing the errors identified using the split-sample test for the main investigation. Surfaces generated using 25% of the data

Algorithm	Resolution	Statistic max	min	mean	std dev	RMSE
Bilinear	0.5m	21.05	-23.98	-0.01	1.39	1.39
	1m	19.09	-23.97	-0.01	1.36	1.36
	2m	20.14	-24.02	0.00	1.35	1.35
Bicubic	0.5m	293.74	-349.76	-0.01	1.93	1.93
	1m	502.56	-352.85	-0.01	1.74	1.74
	2m	50.06	-363.11	-0.01	1.60	1.60
NN	0.5m	23.52	-24.76	-0.02	1.71	1.72
	1m	23.85	-24.76	-0.02	1.73	1.73
	2m	23.85	-24.83	-0.01	1.80	1.80

maximum and minimum error were caused by spiking within the cubic surfaces. The generation of these spikes (both positive and negative) was observed to occur in areas of complex elevation, such as those areas with a high rate of surface change. The positive spikes were much larger in magnitude than the negative dips in surface. The location of the spikes in one of the cubic surfaces is shown in Figures 7.6 to 7.8, these figures show that the spikes in the bicubic surface tend to occur at the edges of features where two points may occur close together in planimetry but represent very different height values. The spikes were therefore deemed to be local artefacts, caused by the absence of a constraint on the maxima and minima predictions to be tied to the original data points. The occurrence of the spikes was likely to be related to their location at the edges of features. In these areas there are rapid changes in elevation, and the larger values for the positive spikes in these areas were thought to be caused by the positive height values of the surface objects causing the cubic algorithm to over-estimate the vertical limit of the step change.

There was also a possibility that the spikes could have been caused by the slight differences in point density that can occur at the edges of features in laser datasets, and which is dependent on the angle of the sensor. However, if this were the case then problems with the other interpolators should also have been noted.

The magnitude of the spikes was a cause for concern, and suggested that this interpolation method may not be suitable for modelling complex elevation areas. This feature of the cubic algorithm had not been noted in the pilot study, possibly because of the small size of the study region used in the preliminary investigation. However, the repeat occurrence of the spikes for all of the experiments run for the main investigation perhaps suggests that this algorithm should not be used without some additional filtering to remove these erroneous surface characteristics. Any 'spike filtering' algorithm would increase the processing times for the algorithm, and this may be an important consideration for the user. Finally, the nearest neighbour algorithm produced slightly higher errors than the bilinear, which was consistent with the findings of the pilot study.

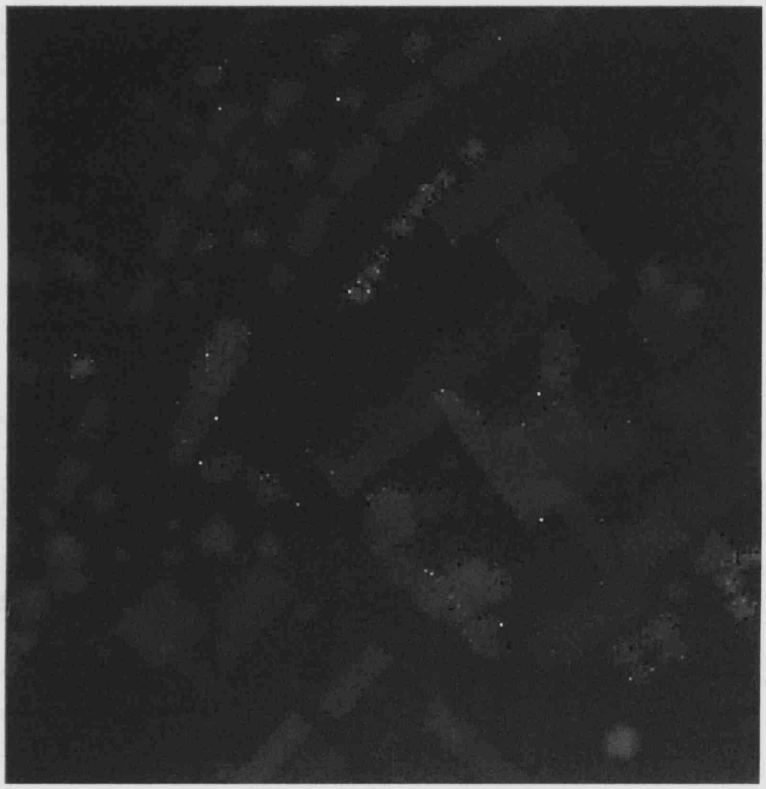

Figure 7.6: Example of area dominated by spikes in cubic surface. The locations of spikes can be seen by the 'freckling' of the surface. Darker pixels indicate negative spikes, whilst the lighter 'freckles' indicate positive elevation spikes. Extent of area shown is approximately 225m by 235m

Figure 7.7: Example of area dominated by spikes in cubic surface. Extent of area shown is approximately 225m by 235m

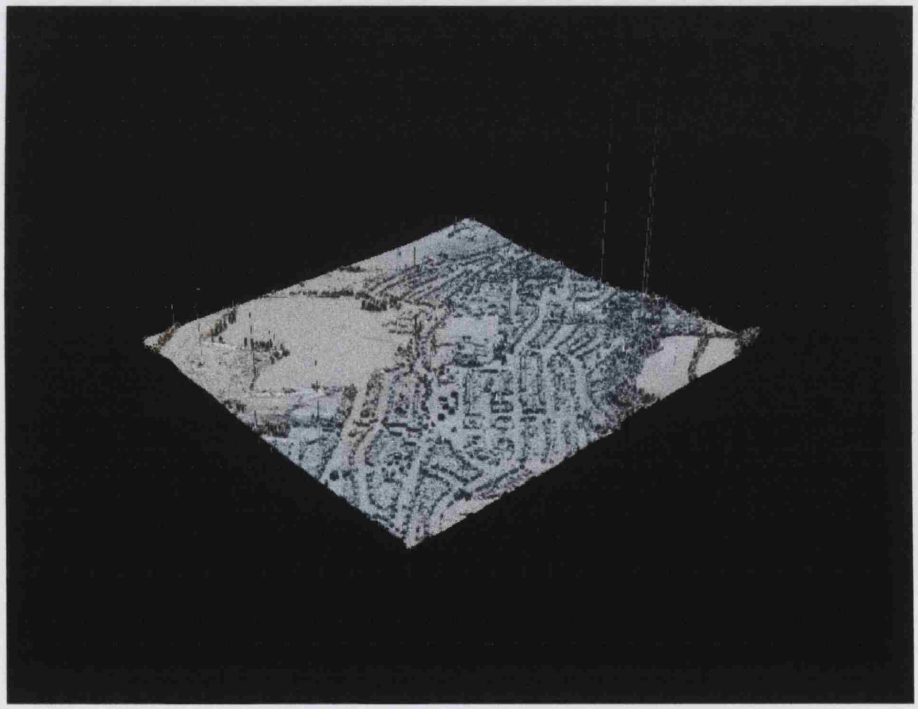

Figure 7.8: View of the surface created using the bicubic algorithm at 1m resolution. Note the spikes across the surface.

In terms of the stability of the algorithms, comparison of Tables 7.1 to 7.3 shows that as the number of points used for surface generation decreases, error increases. Whilst some increase in estimated error was expected, it was noted that these were very small. This was slightly surprising given that as the data become more sparse, so the assumptions made during the interpolation were expected to make differences between methods more apparent. In practical terms, this means that if the point density varies across the study area then the user's choice of method and parameters can be critical. However, these results would suggest that the behaviour of the methods was reasonably predictable as the estimates of the error did not change greatly in the three experiments. In the case of the bilinear algorithm there was only an increase of between 1 and 3cm, and even for the relatively unstable bicubic method the differences were generally small. The only major differences between the statistics were the results of the spike and dip artefacts in the bicubic surfaces.

Finally, the characteristics of the surface estimated errors were found to change as grid resolution was altered. The changes were not the same for all of the methods. For the bilinear algorithm there was generally a reduction in estimated error for the 1m grid spacing. This was the resolution which was the closest to the original point density, and therefore these results confirmed those from the pilot study. The nearest neighbour algorithm produced very similar results for the 0.5m and 1m resolutions, with RMSE values only differing by 1cm. The bicubic results were again largely influenced by the spikes within the surface estimates. The magnitudes of the spikes decreased as the spatial resolution decreased, and for this reason the 2m spatial resolution produced the lowest error estimates for this algorithm. These results were not thought to be reliable simply because of the presence of the spikes. The magnitudes of the spikes and dips in the 0.5m surfaces were larger than for any other grid resolution. This may have arisen because the bicubic algorithm is continuous in the first derivative, in other words it attempts to produce a smooth surface. The interpolation between the higher density grid points simply serves to increase the likelihood of spikes occurring, and because the increased number of spikes must be accommodated within a smooth surface and this can create even larger spikes. The

magnitude of these spikes was nevertheless still surprising, and certainly for use in urban elevation modelling was concerning.

The distribution of the calculated residuals was found to be a leptokurtic normal distribution, as shown in Figures 7.9 to 7.11: showing histograms of the residual values plotted into one hundred equal bins. The fact that the residuals are approximately normally distributed suggests that the assumptions discussed in Chapter 5 regarding the probable range of error values hold true.

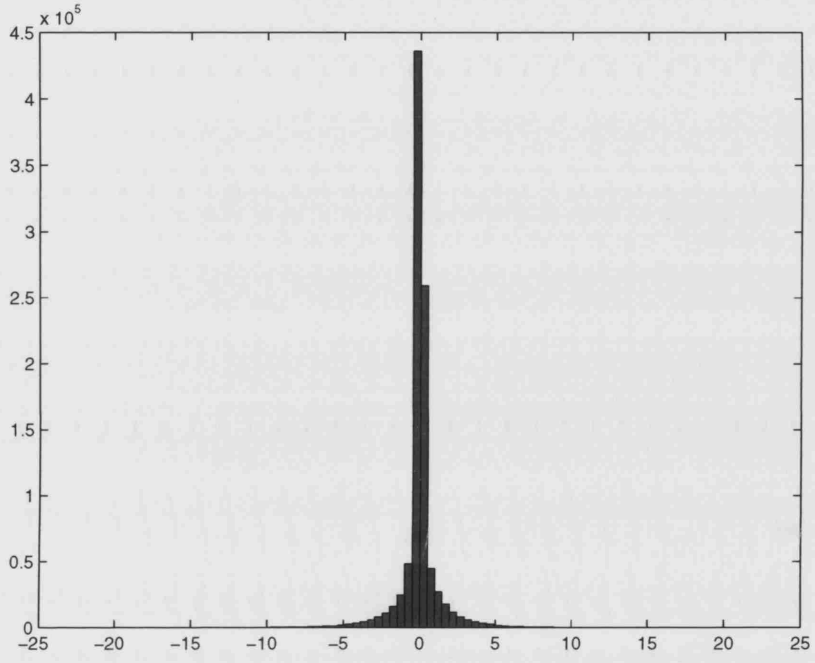

Figure 7.9: Histogram showing the distribution of residual values for the bilinear 1m surface. Histogram shows all values, with no outliers removed

In order to qualitatively assess the spatial pattern of errors, scatter plots were created. These showed the location and magnitude of errors for each of the surfaces generated. Figure 7.12 (page 209) shows all of the errors for the bilinear 0.5m resolution DSM. It is difficult to differentiate between points when plotted *en masse* in this way. To show the patterns more clearly, a number of thresholds were used to show the locations of some of the higher magnitude errors (Figures 7.13 to 7.15). As with the pilot study findings, it appeared that many of the higher errors occurred

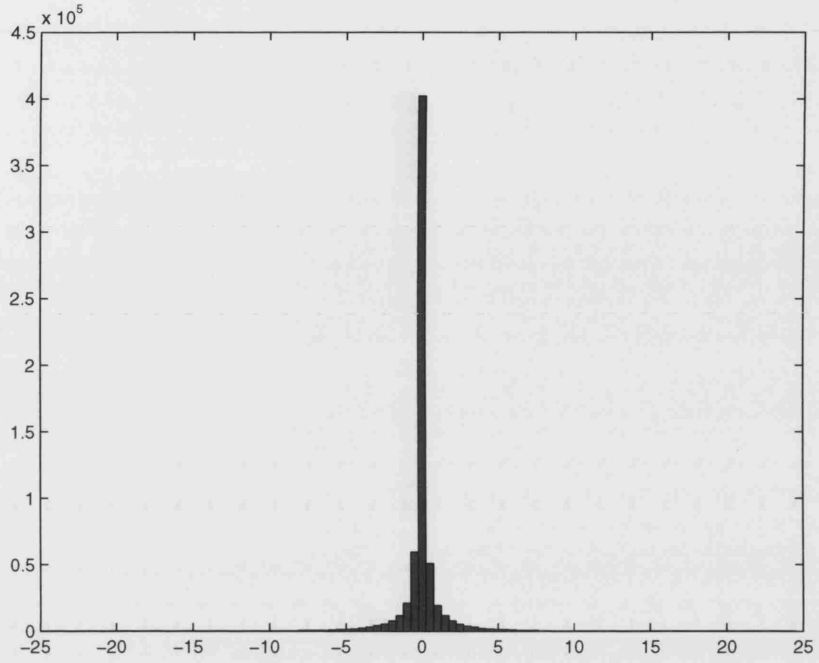

Figure 7.10: Histogram showing the distribution of residual values for the bicubic 1m surface. The histogram shows the distrubution of residuals between -25 and +25m, this excludes the outliers. The full range of residuals for the bicubic surface was between -431.02m and 931.72m, with 165 outliers below the -25m threshold and some 128 outliers above +25m

across object boundaries, and especially over vegetation. In order to quantify this observation a number of approaches were adopted.

Firstly, the intensity of the green band of an aerial photograph of the study area was calculated to provide an approximate vegetation classification. Given the observation that many of the largest errors occurred over vegetation, it was considered that there may be a correlation between the magnitude of error and the intensity of the green band in the image. The experiment was based on the work of Itti et al. (1998), this method was converted to MatLab code for use in this investigation by Dr I. Sargent (Ordnance Survey). The results of this are shown in Figures 7.16 and 7.17. Unfortunately, the results of the intensity calculations showed that it was difficult to distinguish between grassed bare earth areas and wooded regions on the basis of green intensity values, and as such this was an unsuitable method for determining whether there was a relationship between the higher magnitude errors

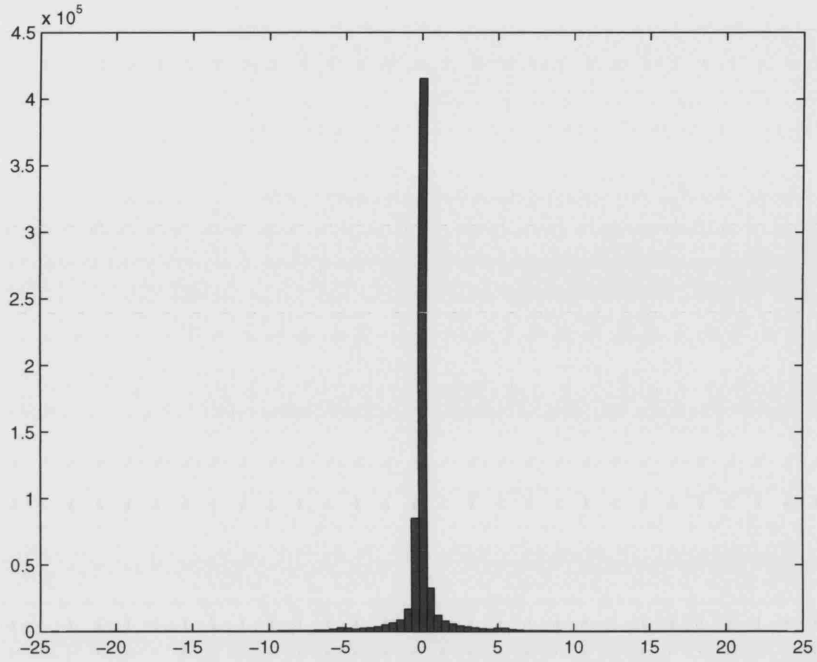

Figure 7.11: Histogram showing the distribution of residual values for the nearest neighbour 1m surface. Histogram shows the distribution of residuals between -25m and +25m, excluding two outliers, one which lay at -25.42m and one which lay at 25.68m

and the incidence of vegetation.

As an alternative to the intensity approach, an aerial photograph of the study area was classified and the highest magnitude errors manually compared with the classes to determine if there was a relationship between the incidence of error and the vegetation class in particular. Clearly, there are many alternative classification approaches and different algorithms will produce different classifications which would, in turn, affect the results of any comparison. It is beyond the scope of this investigation to assess the relative merits of different classifiers here, but the object oriented approach used here was deemed to be most appropriate given that the only imagery of the study area available was an RGB aerial photograph (see Figure 7.1). In the classification approach used here the image was first segmented into regions, specific values for each of the regions were then calculated. Computed values included mean brightness of the pixels within the region. These values were then used to develop

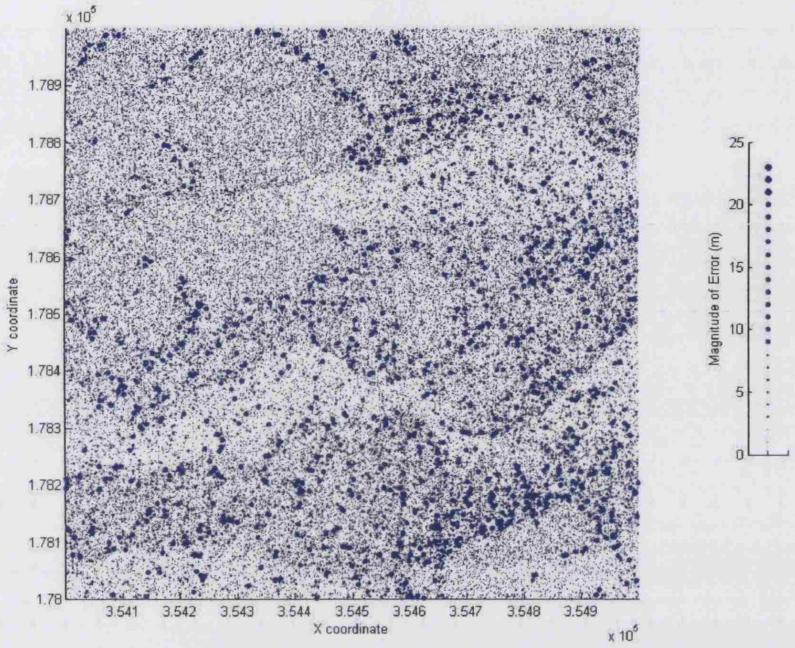

Figure 7.12: Scatter plot showing location and magnitude of errors in bilinear 0.5m surface for the whole Bristol 1km^2 study area

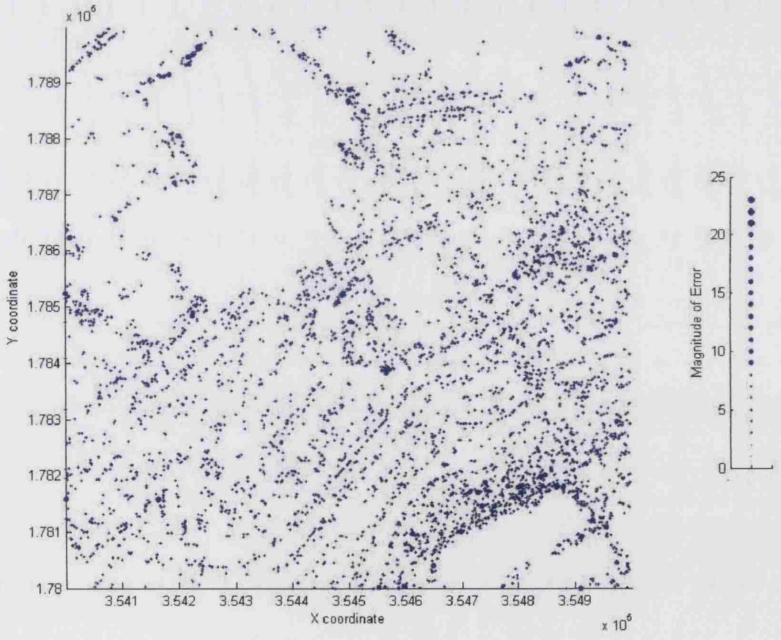

Figure 7.13: Scatter plot showing location and magnitude of errors over 2m in bilinear 0.5m surface

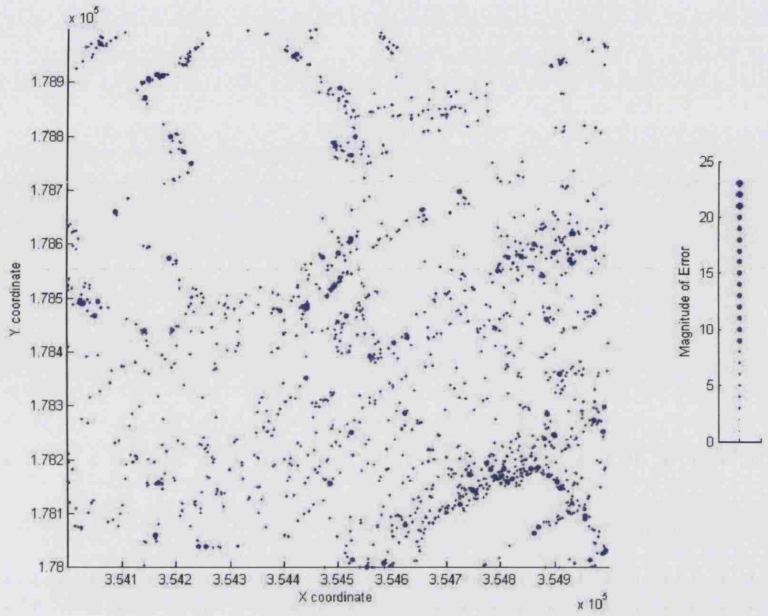

Figure 7.14: Scatter plot showing location and magnitude of errors over 4m in bilinear 0.5m surface

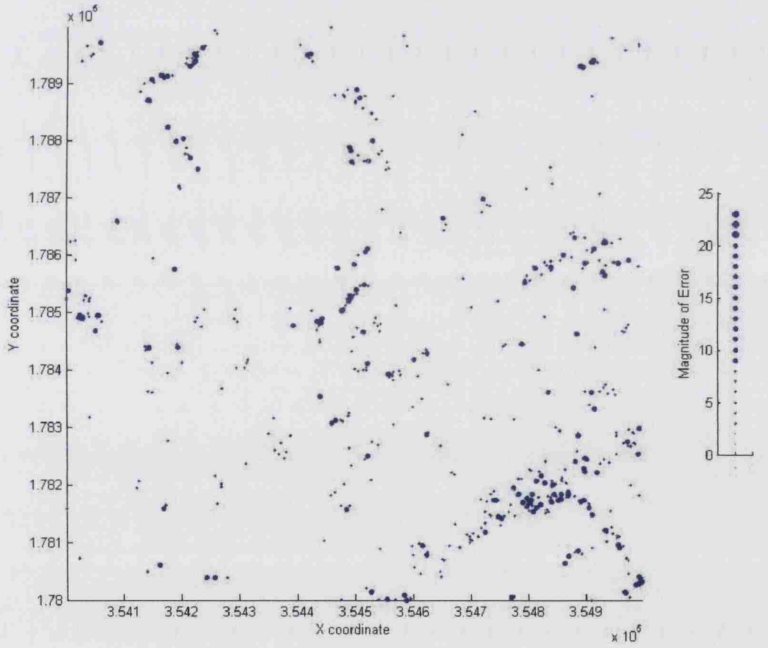

Figure 7.15: Scatter plot showing location and magnitude of errors over 6m in bilinear 0.5m surface

a profile of characteristics for each of the three classifications required (vegetation, bare earth and buildings), and the classification performed based on these different

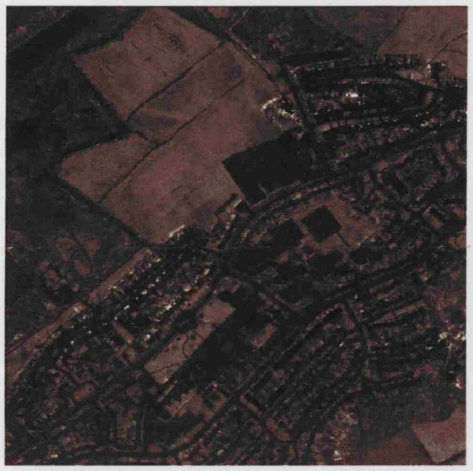

Figure 7.16: Showing the intensity of the green band, displayed in greyscale

Figure 7.17: Showing the intensity of green band of original image. Scale bar shows green intensity

profiles. The resultant classification is shown in Figure 7.18. The accuracy of the classification was assessed manually and is shown in Table 7.4. Of most importance to this study is the producer's accuracy assessment for the vegetation class, which as can be seen from table 7.4 was 100% indicating that all of the pixels classified as vegetation were confirmed as being vegetation (by manual comparison with the aerial photograph).

The locations of the highest magnitude errors were then compared with the three classes to determine if there was a predominant class within which the greatest errors occurred. For this investigation only those errors greater than 6m were used. At a confidence level of 95% and confidence interval of 10, 81 sample points of the 499 errors above 6m were required. For a randomly selected sample, 63% of estimated errors were found to coincide with pixels classed as vegetation, whilst 30% occurred over bare earth and some 7% over buildings. Clearly, whilst there are some limitations of this investigation, specifically with regard to its classification accuracy and more importantly that a random smapling strategy was employed which means that the values may be proportional to the presence of each category in the image. This said, the results do at least suggest that there is some relationship between the magnitude of error and the incidence of vegetation. If this could be shown to be the case, further research might focus upon the analysis of error patterns as an aid

to classification, or to the processes of filtering and segmentation. This is discussed further in Chapter 11. Further research should investigate this potential relationship further by using a stratified sampling strategy in the first instance to compare the incidence of errors with vegetation.

Figure 7.18: Resultant classification of aerial photograph

7.2.2 Analysing the Height Errors

Height errors in the generated surfaces were investigated by comparing the modelled heights with those in the terrestrial laser data. Points for comparison were made close to the site of terrestrial laser data capture, so as to reduce the effect of the systematic errors which can sometimes arise in terrestrial laser scanning. Differences between the interpolated surfaces and the terrestrial laser points were recorded and descriptive statistics calculated: these are presented in Table 7.5. It was found that

in terms of mean error, the bilinear surface contained the least error while the nearest neighbour surface had a much higher mean error value of some 1.27m. The RMSE values for each of the methods were found to be similar - although again the nearest neighbour was found to produce higher errors. The RMSEs ranged between 2.21 and 2.79m for the different methods. From the results in Table 7.5 it can be concluded that the restructuring process for this experiment introduced at least 2.08m of error. This is a substantial amount of error, and may have serious repercussions in terms of the quality of analyses which may subsequently be made using the surfaces, including post-processes such as filtering and segmentation. These issues are discussed further in Chapters 8 and 9 respectively.

7.2.3 Summary of Findings from Main Investigation

The results from the main investigation largely corroborated the findings for the pilot study. In particular, the bilinear interpolation algorithm was found to produce the lowest estimate of error, and for this algorithm the optimal grid spacing was also found to be that closest to the original point density. The nearest neighbour was again found to produce slightly higher estimated error than the bilinear method. The main difference in these results compared to those of the pilot study was for the bicubic method which was found to produce very large spikes and dips. For this reason was thought not to be a suitable method for modelling urban elevation surfaces for this density of laser points without a further processing stage to filter the erroneous spikes from the surface.

7.3 Significance of the Findings for Urban Flood Modelling and Other Real World Applications

The relevance of these findings can be exemplified with reference to flood modelling, which is a common application of ALS DEMs. The modelling of flood predictions

Table 7.4: Contingency Table Showing Accuracy of Aerial Photograph Classification

		Truth Veg	Building	Bare earth	row total
Classification	Veg	43	4	7	54
	Building	0	15	8	23
	Bare earth	0	0	27	27
	column total	43	19	42	81.73%
Producer's Accuracy					
	veg	100.00%			
	building	78.95%			
	bare earth	64.29%			
User's Acuracy					
	veg	79.63%			
	building	65.22%			
	bare earth	100.00%			

Table 7.5: Descriptive Statistics from Validation Investigation

Dataset	**Statistic**				
	Min error (m)	Max (m)	Mean (m)	Std Dev(m)	RMSE(m)
Bilinear 2m	-2.09	6.74	0.66	2.48	2.56
Bilinear 1m	-7.74	5.43	0.18	2.34	2.35
Bilinear 0.5m	-8.15	5.62	0.46	2.40	2.45
Bicubic 2m	-1.47	5.98	0.75	2.08	2.21
Bicubic 1m	-7.78	5.35	0.56	2.45	2.51
Bicubic 0.5m	-8.12	5.61	0.61	2.52	2.59
NN 2m	-5.58	6.62	0.78	2.40	2.53
NN 1m	-5.10	7.53	0.98	2.61	2.79
NN 0.5m	-0.92	6.74	1.27	2.23	2.56

is of great importance for disaster mitigation, urban planning and the insurance industry alike. The prediction of floods requires the input of high quality height models, in order to anticipate the extent and severity of an event. Clearly, the predictions of a flood model will be sensitive to the interpolation functions used to create the height model. Small differences in values in the DEM which is used as an input to the flood modelling program can have a large effect on the predictions. The effects of these differences for the surfaces created in the main investigation are shown in Figures 7.19 to 7.26. At this scale, the difference between the flood predictions for the various surfaces are subtle, but they nevertheless remain important given the applications for such models. In Figures 7.19 to 7.26 the terrain is coloured brown/beige. Note especially the differences in flood level in the top right of the study area where the relief changes more rapidly.

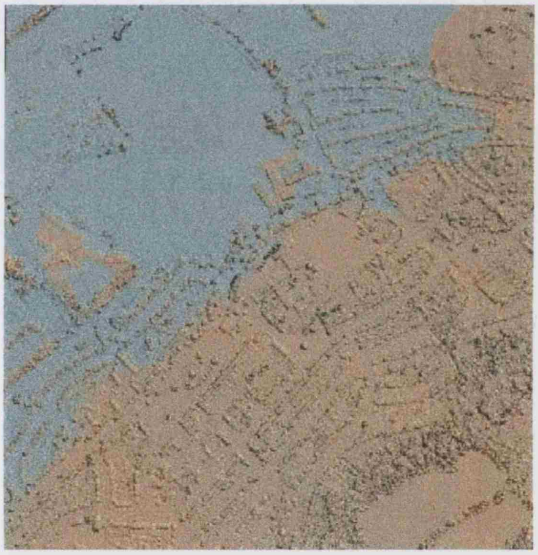

Figure 7.19: Flooding of cubic 0.5m surface with water level of 12m

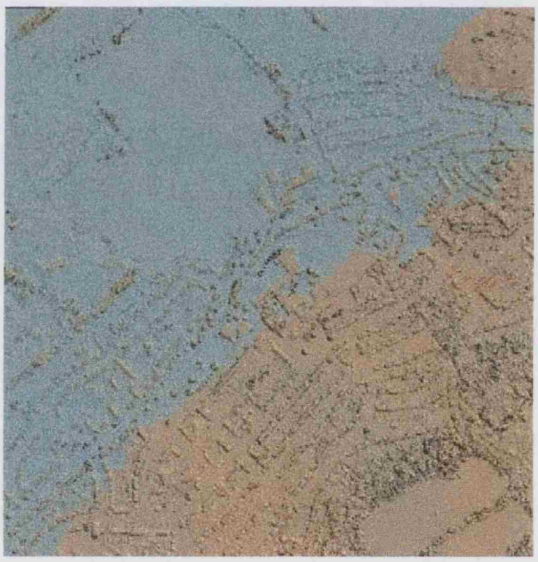

Figure 7.20: Flooding of cubic 0.5m surface with water level of 16m

In order to assess the effect of the different surface generation routines at a larger scale the surfaces created in the pilot study were also used in the flood modelling. The results are shown in Figures 7.27 to 7.36, in which the five surfaces created are 'flooded' to depths of 4m and 5m. Here the terrain is displayed in green to highlight the differences between models at this large scale. These figures show the marked difference between the flood predictions, particularly the limited flooding in

Figure 7.21: Flooding of cubic 0.5m surface with water level of 20m

Figure 7.22: Flooding of cubic 0.5m surface with water level of 24m

Figure 7.23: Flooding of nearest neighbour 0.5m surface with water level of 12m

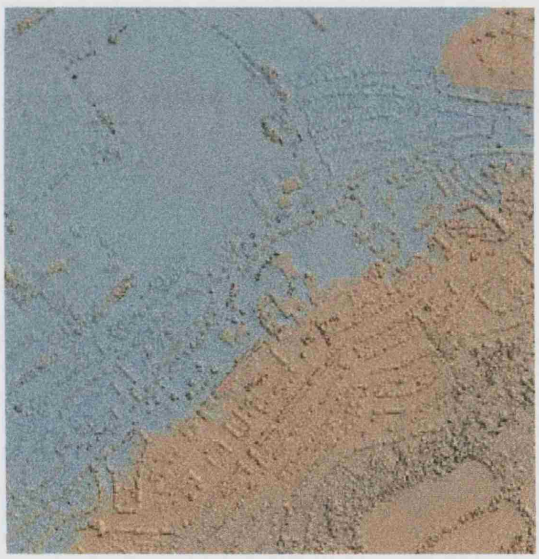

Figure 7.24: Flooding of nearest neighbour 0.5m surface with water level of 16m

the nearest neighbour surface. This is caused by the 'stepped', or blocky, nature of this surface. In the nearest neighbour flood prediction it is clear that the level of the 'step' must be exceeded in order for the flood to propagate. In this instance the stepped characteristic of this surface acts as a flood break, altering the results, and producing a potentially erroneous prediction. In contrast, the biharmonic splined surface, which was observed to produce a very smooth surface, generated the most extensive flood prediction. Clearly, if these models were to be used in the calculation and mapping of flood risk areas, the use of different gridding techniques could substantially alter the results.

The same argument holds for the use of DEMs in similarly sensitive applications, such as mobile phone wave propagation modelling, and noise pollution modelling. Thus differences in the height values of DEMs can have significant implications where the models are used in a predictive or analytical capacity. Laser scanning height models are already being used in this analytical context (e.g. Schofield et al. 2003), the immediate need for information about the potential differences in models is evident.

Figure 7.25: Flooding of nearest neighbour 0.5m surface with water level of 20m

Figure 7.26: Flooding of nearest neighbour 0.5m surface with water level of 24m

Figure 7.27: 4m Flood Event Across Bilinear 1m Surface

Figure 7.28: 5m Flood Event Across Bilinear 1m Surface

Figure 7.29: 4m Flood Event Across Bicubic 1m Surface

Figure 7.30: 5m Flood Event Across Bicubic 1m Surface

Figure 7.31: 4m Flood Event Across Nearest Neighbour 1m Surface

Figure 7.32: 5m Flood Event Across Nearest Neighbour 1m Surface

Figure 7.33: 4m Flood Event Across Biharmonic Spline 1m Surface

Figure 7.34: 5m Flood Event Across Biharmonic Spline 1m Surface

Figure 7.35: 4m Flood Event Across Ordinary Kriging Spline 1m Surface

Figure 7.36: 5m Flood Event Across Biharmonic Spline 1m Surface

7.4 Conclusions

The introduction of errors in the gridding of data remains only one source of error in the process of modelling from ALS data. It has been the purpose of this chapter to demonstrate how these gridding errors may be introduced in urban surfaces, and to assess how the magnitude and the spatial structure of these can change with different techniques and scales of analysis.

In terms of the interpolation methods investigated here, it was found that the nearest neighbour algorithm produced the highest mean errors, while the bicubic method produced large oscillations at the edges of some features causing spikes within the resultant surface. These spikes could be eliminated by additional processing, although this may be costly in terms of time for the ALS data user. The bilinear method was found to introduce the least error into the surface. For users wishing to reduce error incorporated during data restructuring the results of this study suggest that the bilinear algorithm should be used. These results differ slightly from those of Rees (2000) who suggested that either the bilinear or the bicubic algorithm would be suitable. However, Rees' (2000) conclusions were based on modelling elevation in non-urban areas, and it is likely that the complexity which characterises urban elevation renders the bicubic algorithm unsuitable for many applications. Furthermore, the results of the pilot study suggested that neither the biharmonic spline or block kriging were suitable for modelling urban elevations.

The grid spacing which introduced the least error was that closest to the original point spacing, this was the case for both the pilot study and the main investigation. This corroborates the findings of Behan (2000).

Whilst differences have been highlighted between different interpolation methods in this chapter, it should be noted that there is no generic optimal creation methodology, as the final decision regarding interpolation algorithm, grid spacing, filtering method, and segmentation procedure must be driven by the requirements of the ap-

plication for which the DEM is intended. It is suggested here that to make informed decisions regarding the specific modelling processes, users should be provided with error information which may come in the form of a map of the spatial structure of error. The provision of this error information further requires that there are software processes which are able to cope with the communication of error, and with the incorporation of this information in subsequent analysis. This in turn requires that the data users understand the error information, and can use this intelligently in order to reduce the introduction of error in their ALS data processing.

The investigation presented here has looked only at assessing the errors in the height models. However there have been a number of studies which have assessed both accuracy and shape preservation of the surfaces. Similar studies could also be applied to the urban environment, and would be particularly useful for applications such as feature extraction where the choice of interpolation method may depend on both accuracy and the preservation of key features.

7.5 Chapter Summary

The effects of a variety of ALS data restructuring techniques are analysed and results presented. Findings suggest that the bilinear algorithm produces the least error of the interpolation methods investigated, and that the grid resolution should be as close as possible to the original point spacing in order to further minimise the introduction of errors. The amount of error found in all of the surfaces was greater than had been expected, and in all cases was greater than 1m. Given that such large errors are introduced during this stage of post-processing, it is evident that users must make informed decisions in order that this error is minimised as far as possible. Finally, the spatial pattern of the highest magnitude errors appeared to coincide with the edges of features within the study areas, with higher errors appearing to occur over vegetation.

The results from this chapter are of import in the following two investigations which analyse the errors in subsequent post-processing. The results of these further investigations are presented in Chapters 8 and 9.

Chapter 8

Analysis of the Error Arising Through Filtering

8.1 Chapter Introduction

It was shown in Chapter 2 that one of the principal stages in laser data post-processing is that of filtering. This is the process of classifying laser points (or pixels) into either bare earth or above ground classes in order to produce filtered datasets containing ground points only. These filtered datasets are then most often used to create Digital Terrain Models (DTMs) for applications such as environmental analysis, urban planning and flood modelling.

Users are faced with an increasingly wide range of filtering algorithms from which they must choose the most appropriate for their post-processing needs. The algorithms work in markedly different ways, using a variety of input data structures and processing techniques, and consequently may produce very different filtering results. However, little is known about the sensitivities of the algorithms or about the quality of the datasets they produce. Some comparisons of a selection of algorithms have been offered in the past (Sithole and Vosselman 2003; Tao and Hu 2001; Huising

and Pereira 1998b; Haugerud and Harding 2001). The general observations from such studies are useful, however in order for users to make informed decisions, and either manage or reduce error, they require more detailed information. In particular, users need to understand how the following factors may affect the accuracy of the filtering:

- input data structure
- spatial resolution
- choice of filtering algorithm
- filtering parameters

The investigation presented in this chapter offers a thorough, robust, quantitative analysis of the effects of these parameters on both the amount and the spatial pattern of errors. In terms of quantifying the post-processing error budget (see Equation 2.13 which was first introduced in Chapter 2, this chapter presents the methodology for quantifying filtering errors (ε_f).

8.2 Chapter Aims

In this chapter, three different filtering algorithms are used to create a variety of bare terrain datasets. The inputs to these algorithms are firstly, the gridded datasets at three spatial resolutions created in the restructuring investigation (see Chapter 6), and secondly the raw laser data.

The aim of the investigation is to establish whether using different input data, or different filtering algorithms has an effect on the amount of error in the resultant bare earth dataset in urban applications. If it is found that there is no significant difference between the resultant filtered datasets then the user's choice of filtering method, data structure, or spatial resolution will be of little or no consequence for

the amount of error introduced at this stage. However, if the converse is found, then the user's decisions may have important consequences for model quality, and the results of the investigation will highlight where emphasis needs to be placed in the post-processing design.

In order to structure the investigation, the methodology was designed to specifically answer the following research questions:

- How does the choice of data structure affect the results of the filtering?
- Does the choice of gridding method affect the characteristics of error in subsequent filtering?
- Does the choice of filtering algorithm make a significant difference to the accuracy of the filtering?
- How sensitive are the filtering methods to changes in parameter settings?
- Does the spatial resolution of the data affect the characteristics of the error in the filtered datasets?

The chapter is divided into four sections. The first introduces the filtering algorithms used; the second presents the method for quantifying the success of the methods; the third section presents the results of the comparisons including the answers to the specific research questions; and the fourth discusses the significance of the findings.

8.3 The Filtering Algorithms

The three algorithms used in this investigation work in very different ways, and require detailed explanation. The grid based geometric algorithm is original and was written for this investigation, whilst the Iterative Robust Interpolation (IRI) and Adaptive TIN Method (ATM) (both of which use discrete data) are published

methods which have been specifically adapted for this investigation. The details of each of the three algorithms are presented below.

8.3.1 The Grid Based Geometric Classification Filtering Technique

The geometric filtering algorithm works with gridded laser data which are held as a raster, and in this way can exploit many pre-existing image processing techniques. The use of grid-based algorithms is reasonably common, and examples include Brovelli et al. (2002), Elmqvist (2001), Elmqvist et al. (2001), and Wack and Wimmer (2002). However, it is sometimes assumed that grid-based algorithms can introduce error into the dataset in the data gridding process and that this error is then exacerbated during the filtering process. One of the aims of this study is to assess critically the validity of this long-held assumption.

There are a number of stages in the geometric algorithm, which are summarised here, before being explored in more detail below. In the first stage, the laser image is segmented by identifying clusters of pixels with similar characteristics, using E-Cognition software (Definiens). These clusters are subsequently grouped into meaningful objects in a process called *image division.* This method represents an object-oriented approach to modelling which differs from the vast majority of standard image processing techniques, which tend to only use information held by a single pixel. The approach adopted here works on the assumption that important semantic information is lost in the processing of individual pixels, and that much more meaning can be derived from the identification of image objects and their mutual relationships. An example of the image division process is explored below.

Figure 8.1 shows one of the original laser data images which is to be segmented. At this stage of processing a number of parameters for image division are established in order to define the characteristics of the objects to be identified. Parameters include the importance of the shape and compactness of the identified objects. The default, or generic, settings which were used for the algorithm are shown in the first column

of Table 8.1. To test for the sensitivity of the algorithm these parameters were later changed to those shown in column two of Table 8.1, as these were thought to be more suited to the study dataset. In the example shown here the default parameters were used to divide the laser into groups of pixels (Figure 8.2). These divisions are shown more clearly in Figures 8.3 and 8.4.

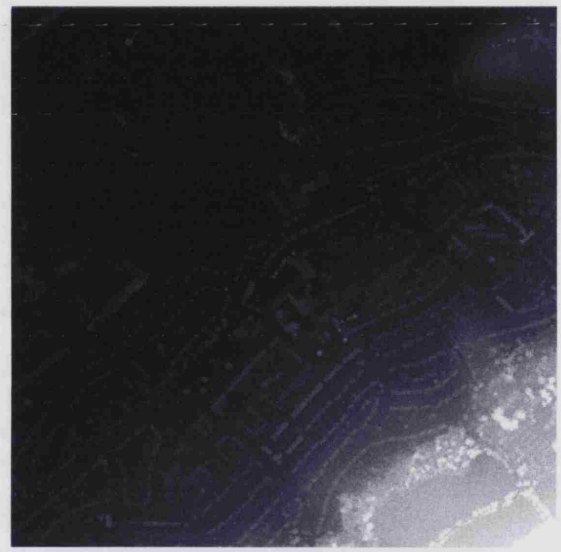

Figure 8.1: Greyscale height image of the bilinear 2m surface before image division for the whole Bristol study area

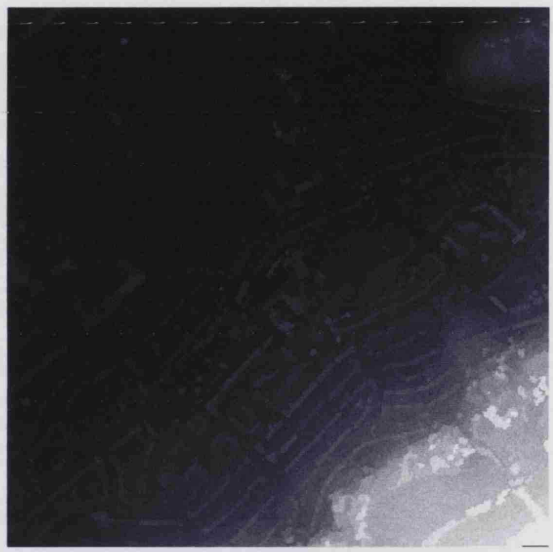

Figure 8.2: Showing the divisions of the image created from default parameter settings for the whole Bristol study area

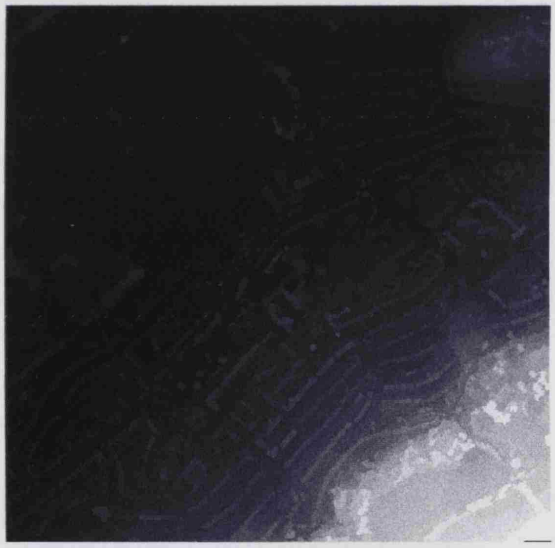

Figure 8.3: Showing the outline of one of the identified objects for the whole Bristol study area

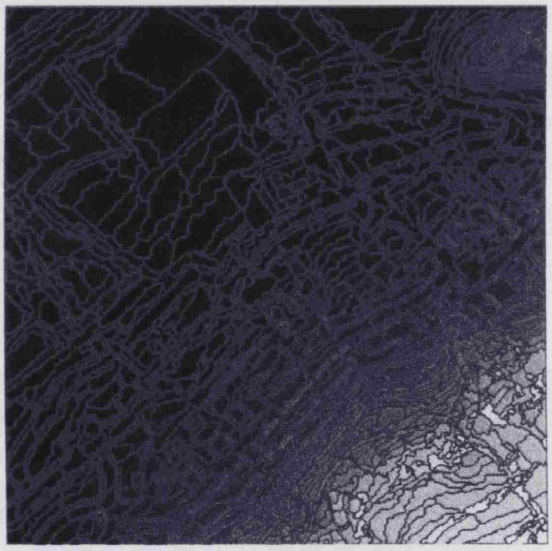

Figure 8.4: Showing all of the image divisions in the whole of the Bristol study region

Some explanation of the parameters defined in Table 8.1 is required here. The scale

Table 8.1: Parameter settings for Image Division

Parameters	Default Settings	Adapted Settings
Scale parameter	10	10
Colour	0.8	1
Shape	0.2	0
Smoothness	0.9	n/a
Compactness	0.1	n/a

parameter simply specifies the size of the image objects (in pixels). The remaining parameters are together called the homogeneity criteria and are used to determine various characteristics of the image objects. Firstly, the most important of the parameter settings for the classification is the colour criterion. In the laser image the colour of any pixel relates to the height above datum, and in this way can provide a great deal of useful information about the classification of the associated image object. Secondly, the shape criterion is used to define the weighting (or importance) of the shape of the image object to its classification. In the adapted setting, shape was not considered to be an important factor for determining the classifications of the laser image. The values for the colour and shape criteria define the relative importance of these two factors in the division of the image into objects. If the shape criterion is attributed a weighting value (as in the default settings in Table 8.1), then two further weighting values must be assigned in order to determine the shape of the image objects. These values relate to the relative importance of the smoothness and compactness of the objects.

The objects within the divided image (Figures 8.3 and 8.4) are subsequently classified, based upon their principal geometric characteristics. These characteristics are calculated by quantifying the relationships between the object polygons and their neighbours - such as the mean difference between one object polygon and all of its neighbours. The object polygon is then assigned the mean value for whichever characteristic has been chosen, and the remaining polygons are assigned values in the same way. The image can then be re-displayed showing the recalculated values for each polygon. The values within this new image can then be used in a classification to identify two classes, bare earth and above ground objects, based on the values

of each polygon. A pilot study was conducted to determine which characteristics should be calculated. Candidates included:

- Object shape

  Potential object shape parameters included calculations of object compactness (the result of the calculations of object compactness are shown in Figure 8.5), object length/width ratio (Figure 8.6), and object width (Figure 8.7). Of these shape parameters the object width calculations seemed to produce the most successful results.

- Mean heights

  This would be the simplest way to classify the objects within the image (the results of which are shown in Figure 8.8). However as the values within the image represent heights above datum rather than the heights of the objects themselves this classification did not produce suitable results.

- Mean height differences

  Here the value of each segment was assigned a ratio value which was calculated by comparing it to each of its neighbours. This produced a much more meaningful result than the simple partitioning based on mean heights (see above). The resultant image from the mean differencing division is shown in Figure 8.9.

The success of the parameter settings was assessed qualitatively by visually comparing the resultant images. The most successful were assumed to be those that produced clear differentiation between known objects and ground (through comparison with the aerial photograph of the study area). In accordance with this criterion it was considered that the calculation of mean differences produced the most successful image division (Figure 8.9). The results from this calculation were used as an input for classification of the whole image. Membership of the classes was defined using a weighting function. Segments representing the bare earth were defined as those whose mean difference value fell in the range -10m and 2m different to neighbouring objects. In a similar way, above ground objects were classed as

Figure 8.5: Showing the results of the calculation of object compactness for each of the image divisions

Figure 8.6: Showing the results of the calculation of the length to width ratio for each image division

Figure 8.7: Results of calculation showing object width values

Figure 8.8: Results of calculation of mean object heights

those segments whose values ranged between 1.6m and 50m. The resultant binary classification is shown in Figure 8.10. To this point, the method for the geometric algorithm was implemented using E-Cognition software (version 3, 2002, Definiens Imaging GmbH).

In the subsequent stage of the algorithm the binary classification (shown in Figure 8.10) was used as a filter to determine which pixel values were to be used in the DTM computation. The binary classification was used to 'cookie-cut' the original image to eliminate the above ground pixels (Figure 8.11). The outer pixels for each resultant hole were then removed in order to eradicate the halo effect around each. These halos had been caused by some uncertainty regarding membership classification at the edges of identified image objects, and were eliminated by passing a pre-defined kernel across the image in which the minimum value within the kernel was attributed to the halo cells only. In effect this dilated the gaps. The halos are shown in Figure 8.12.

The resultant gaps are then interpolated across to create a continuous filtered surface, such as that shown in Figure 8.13 below. The interpolation across the gaps in the surface was achieved by passing a kernel over the gap which returned the mean of the pixels in the focal window around each pixel of gap cells. It is difficult to see some of the detail in the image of the DTM, however Figure 8.14 shows the same image which has undergone a histogram equalisation stretch in which the exact form of the surface is more easily identified.

All of the resultant DTMs created using the geometric filtering algorithm are shown below (Figures 8.15 to 8.31). The scripts for creating the filtering using this method are presented in Appendix E).

In terms of qualitative analysis of the resultant surfaces, the success of the DTMs was assessed in terms of how many 'above ground' features appeared to remain in the dataset. In accordance with this criterion, the nearest neighbour input data appeared to produce the most successful DTMs, whilst the bilinear and bicubic data

Figure 8.9: Results of calculation showing mean difference between objects

Figure 8.10: Resultant binary classification, with above surface objects shown in blue

Figure 8.11: Gaps produced by the removal of above ground objects in the cookie-cutting stage. Note the 'halo' of remaining pixels around gaps

Figure 8.12: Following dilation of the above ground areas the halos have been removed before interpolation over the gaps is performed

Figure 8.13: The continuous DTM created using the bilinear 2m data

Figure 8.14: Histogram equalisation stretch of DTM in Figure 8.13

Figure 8.15: DTM created using default method for bilinear 1m data

Figure 8.16: DTM created using default method for bilinear 0.5m data

produced reasonably similar results. It was noted that as grid spacing was reduced there was an increase in the number of above surface objects remaining in the DTMs, this was most likely a product of the parameter settings for the algorithm, and it was thought that different settings (such as those defining the size of the objects for image division) would produce very different results. Finally, it was noted that the adapted geometric filtering method appeared to be more successful than the default method - this was anticipated given that the adaptations were designed to make the parameter settings more suitable for the study area under investigation. The basic qualitative analysis described here was supplemented with more rigorous quantitative investigations, which are presented in Section 8.4.

Figure 8.17: DTM created using default method for the nearest neighbour (NN) 2m data for the Bristol study area ($1km^2$)

Figure 8.18: DTM created using default method for the NN 1m data for the Bristol study area. Lighter pixels represent higher elevation

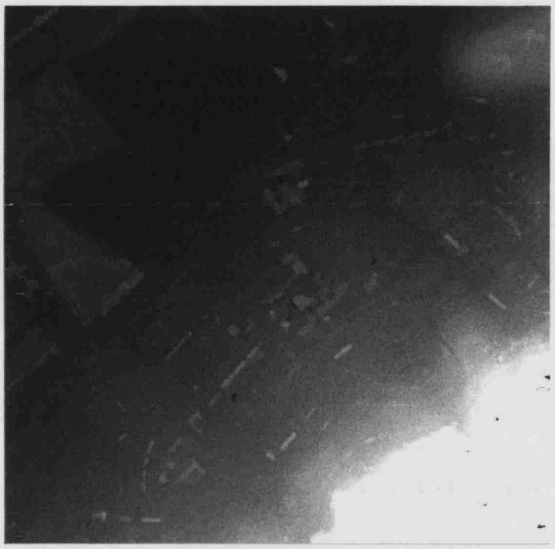

Figure 8.19: DTM created using default method for NN 0.5m data

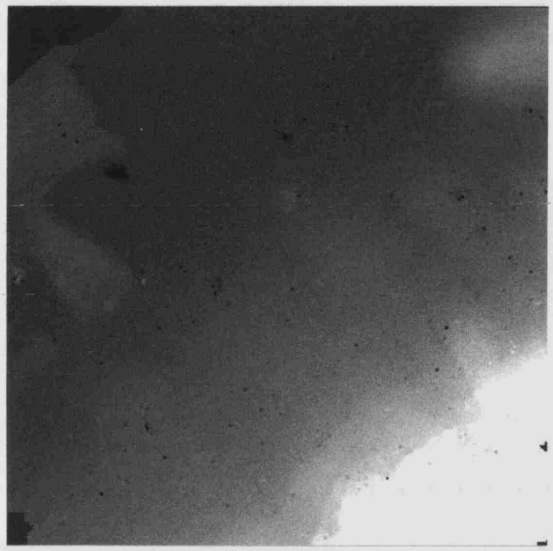

Figure 8.20: DTM created using default method for bicubic 2m data

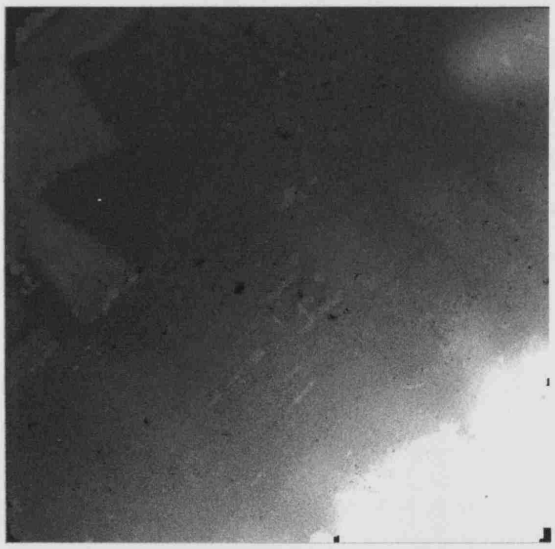

Figure 8.21: DTM created using default method for bicubic 1m data

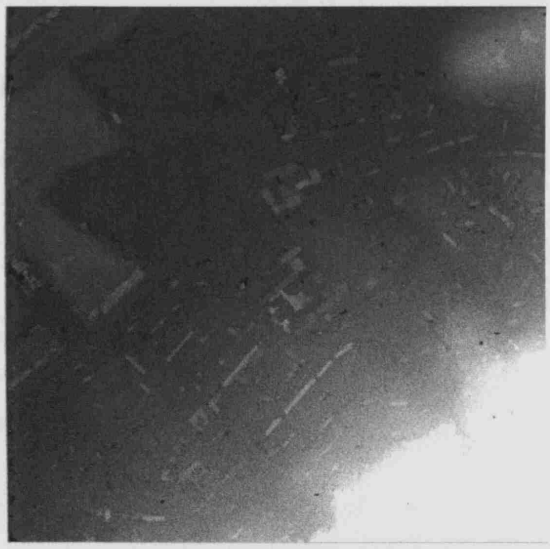

Figure 8.22: DTM created using default method for bicubic 0.5m data

Figure 8.23: DTM created using adapted method for bilinear 2m data

Figure 8.24: DTM created using adapted method for bilinear 1m data

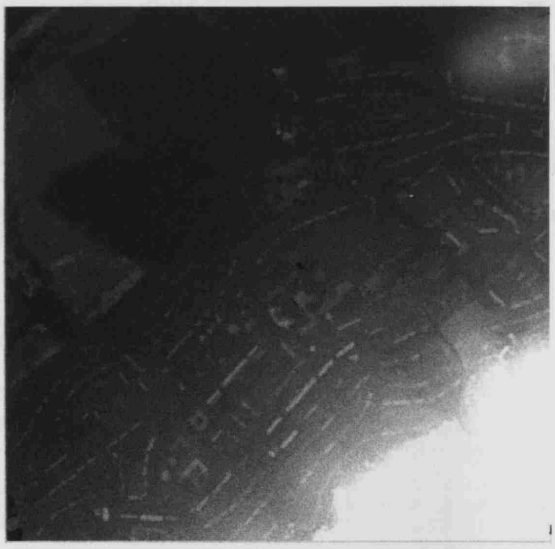

Figure 8.25: DTM created using adapted method for bilinear 0.5m data

Figure 8.26: DTM created using adapted method for NN 2m data

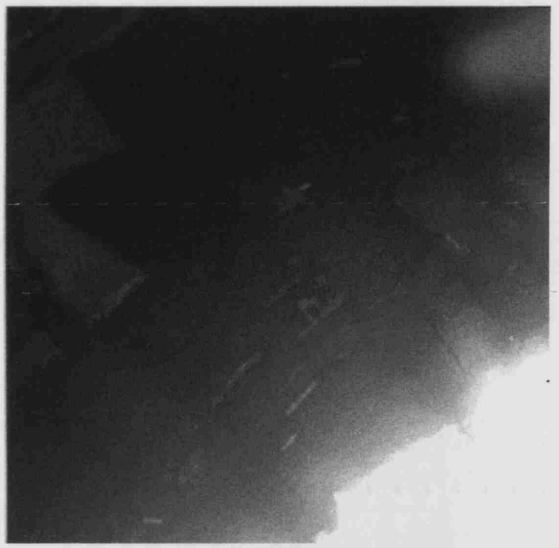

Figure 8.27: DTM created using adapted method for NN 1m data

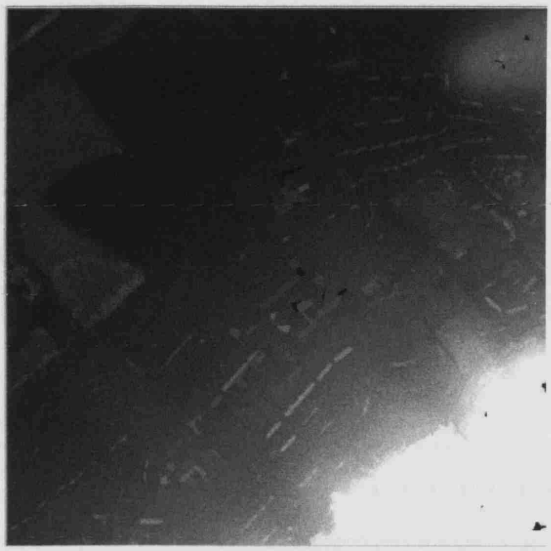

Figure 8.28: DTM created using adapted method for NN 0.5m data

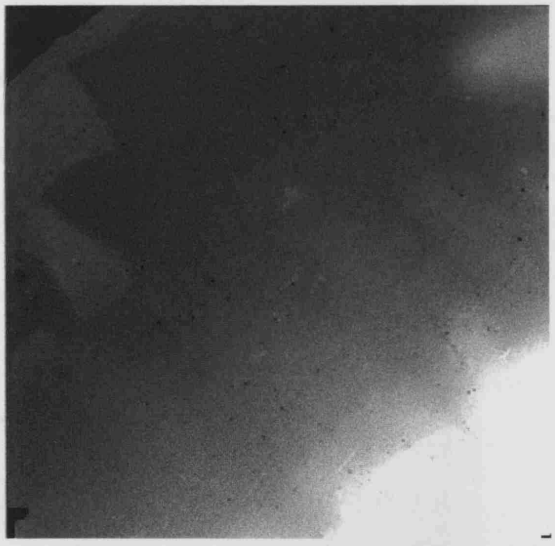

Figure 8.29: DTM created using adapted method for bicubic 2m data

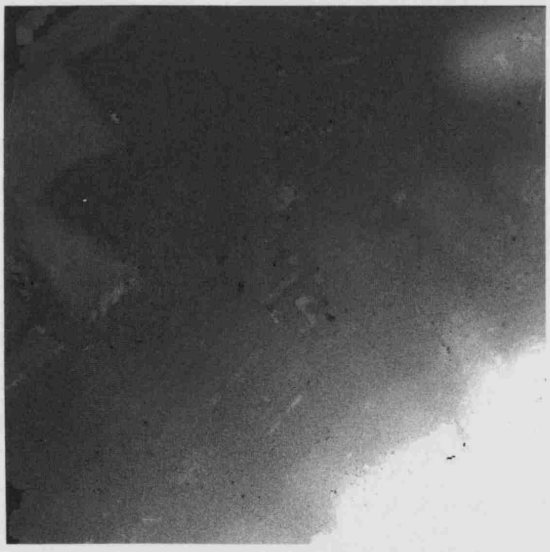

Figure 8.30: DTM created using adapted method for bicubic 1m data

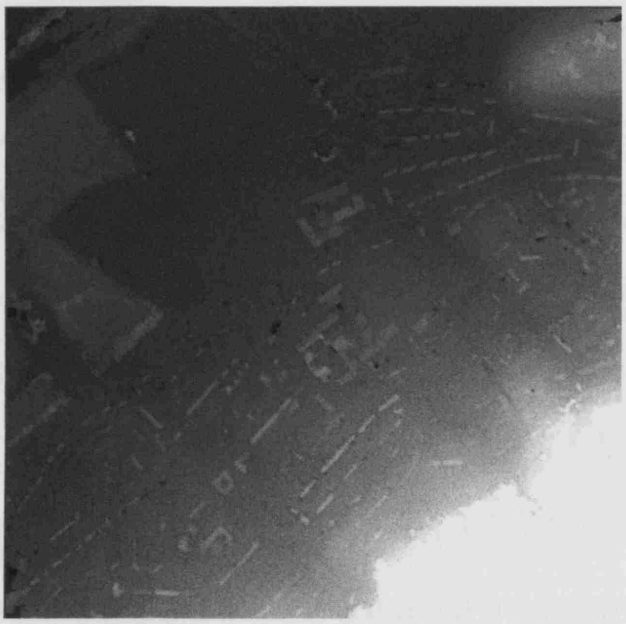

Figure 8.31: DTM created using adapted method for bicubic 0.5m data

8.3.2 Iterative Robust Interpolation (IRI) Filtering

The original methodology for robust interpolation filtering was developed at Vienna University of Technology by Briese et al. (2002) and Kraus and Pfeifer (1998), and has been recently implemented into the SCOP++ software environment (Pfeifer et al. 2001b) . The method developed by the aforementioned researchers was used in this investigation along with an adapted version of the method which was created for this study. The results from the adapted method and from the pre-written algorithm were compared to test the sensitivity of the process to parameter changes.

Hierarchic robust filtering is a bare earth modelling technique, which works with discrete data points and classifies these as either terrain or non-terrain. There are two resultant datasets: a gridded DTM, and a file containing the classified points as a standard (x, y, z) file. The algorithm operates in a hierarchic fashion, creating data pyramids in a coarse to fine strategy in order to accelerate the process. The method first computes an approximation of the surface using only certain points, which are selected based on simple analysis of the raw data. This is done by defining a grid and calculating a representative value for each cell - this representative value may, for example, be the mean or the lowest value in each cell. An approximate surface is then computed using these cell values, against which the original values are compared and residuals calculated. Each z measurement within the original dataset is then assigned a weight according to the residual (distance between the point and the interpolated surface). The surface is then recomputed, this time considering the weights of the points. A point with a high weight will attract the surface, subsequently resulting in a small residual in the following surface computation, whereas a point with a low weighting will have little influence on the run of the surface (Pfeifer et al. 2001a). The bare earth classification is performed during these iterations. If any point falls outside a given threshold distance from the approximated surface, then it is automatically eliminated completely from the surface interpolation. The process of calculating weights for points is continued until either a stable situation arises or the maximum number of iterations is reached.

There are two principal components to the robust interpolation algorithm: the stochastic model (defined by the weight function), and the functional model (surface interpolation). First, in the stochastic element, the weight function is designed to assign high values to those points on the ground, or just below this, whilst assigning low values to any above ground object. The weight function is defined by the shape of a bell curve. The weight function can be asymmetrical so as to adapt to the distribution of errors in relation to the true surface (Briese et al. 2002).

In order to assess the sensitivity of the algorithm to changes in parameter settings, an alternative method (based on the method described above) was written for this investigation. The adapted technique included more filtering stages, so as to eliminate as many non-terrain points as possible. This adaptation was thought to be more suited to the study data than were the generic, default settings. In the filtering stages lower weight values for the function and larger grid cells were used in the thin out stages in order to incorporate potential terrain candidates. For many of the filter stages the eccentricity of the function shape was altered by lowering the penetration rate defined for the laser data. This effectively lowered the interpolated surface, and permitted the inclusion of more below and on-terrain points.

The strategy outlined above used resolutions of 2m, 5m, 2.5m and original data spacing (1m). The resolutions were chosen during a pilot study, and were considered to produce the most effective filtering results, the relative success, or otherwise, of this method is investigated later in this chapter. The stages of the method created here are outlined below.

1. The data is first reduced by extracting the mean value from each 2m grid cell.
2. The stochastic element of the algorithm (the weighting function) is defined. For the first set of weights, this incorporates both above and below terrain points within the permitted weight values. The weights are used to decide which points should be included as terrain points. An intermediate surface is created using the 2m resolution data.

3. The second data pyramid is created, using a cell size of 5m within which the lowest value is selected.

4. Three filtering stages are employed to refine the weighting of points. This again controls the inclusion of points within the terrain surface. An asymmetric weight function is used in order to assign high weights to points on or below the surface. This has the effect of reducing the influence of off-terrain points.

5. The product from the weighting process is used in the interpolation of a 2.5m resolution surface.

6. The raw data points are then compared to this interpolated surface, and the residuals calculated and used to accept those points which fall within a tolerance value (-3 - +1.5m) as terrain.

7. The third data pyramid is created, using a cell size of 2.5m, within which the lowest value is selected.

8. The points classified as terrain are then filtered again to further refine the selection.

9. The surface is then recomputed at a resolution of 1m.

10. The raw data points are then again compared to the 1m interpolated surface and those falling within a tolerance band of -3m - +2m are accepted as terrain points.

11. These points are then filtered twice more to refine the selection.

12. The resultant surface is interpolated at the final 1m resolution.

An example parameter file used for this adapted IRI algorithm is presented in Appendix F.

The difference in the resultant DTMs is shown in Figures 8.32 and 8.33 which, for clarity, show an enlarged section of two of the DTMs. The building footprints and other remnants of above ground features can be seen in the terrain model

produced using the original methodology. Some of these remaining features have been degraded in the terrain model created using the adapted methodology.

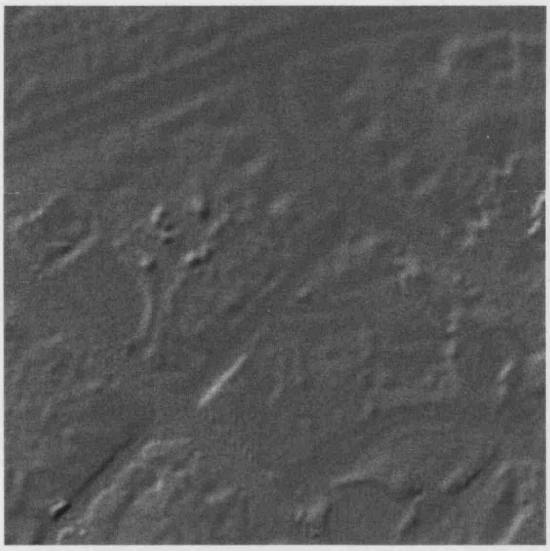

Figure 8.32: Hillshaded image of subset of DTM created using original method

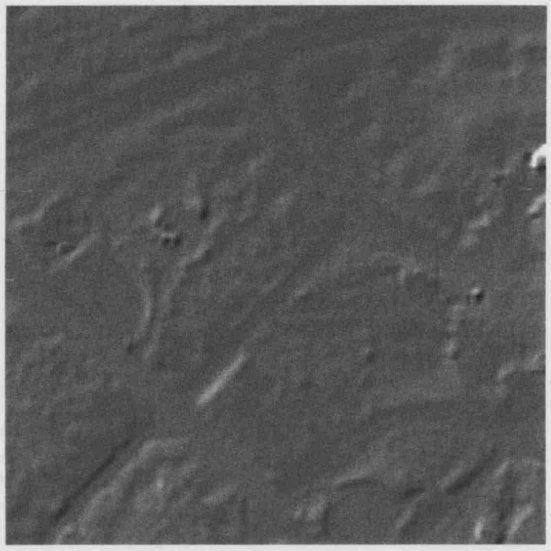

Figure 8.33: Hillshaded subset of DTM created using adapted method

Datasets filtered using the IRI method included the raw data and all of the products from the data restructuring process for each interpolation method at each grid resolution. All of these datasets were filtered using both the original and the adapted methods. In total some 20 DTMs were created for the IRI investigation, all of which are presented in Appendix G: a selection are shown below to highlight some of the principal observations.

Figures 8.34 to 8.37 show the DTMs created from the filtered datasets using the raw data (Figure 8.34), the bilinear data (Figure 8.35), the bicubic (Figure 8.36), and the nearest neighbour (Figure 8.37) input data (at a spatial resolution of 0.5m) with the original IRI method. From visual inspection it is clear that there are a number of differences between the form of the DTMs produced. The results from the gridded DTMs was surprising, with the DTM created from the nearest neighbour gridded data appearing to perform the best. This observation was counter-intuitive given that the error analysis for the data restructuring investigation (Chapter 6) showed that the nearest neighbour algorithm performed the least well of all the investigated procedures since it introduced the highest quantity of error into the

DSM. However, in terms of filtering the nearest neighbour dataset appears to have most of the above ground features removed by filtering (when compared with the original surface shown later in Figure 8.48). Although the nearest neighbour surface appears to be the most successfully filtered, there are still some general similarities between the nearest neighbour, the bilinear and the raw surfaces, with all showing the erosion of many of the above ground features. However, it is clear that the filtering of the bicubic surface has been much less successful, with the majority of features still remaining.

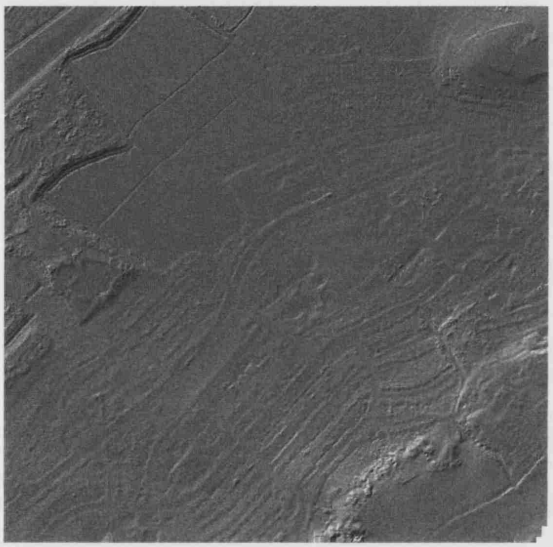

Figure 8.34: DTM created using raw input data (with original point spacing) with the default IRI method

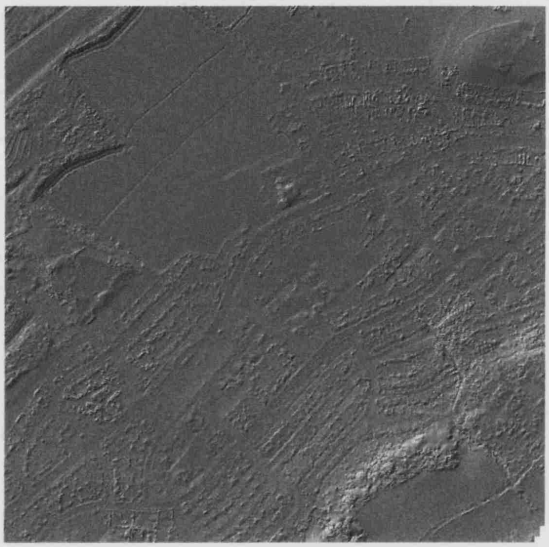

Figure 8.35: DTM created using bilinearly gridded input data at 0.5m resolution with the default IRI method

The resultant surfaces for the adapted IRI method were found to produce visually similar results to those from the original method, with one major difference being that in most cases the original filtering method appears to have removed more of the larger above ground objects and in this respect seems more successful.

The observations from this qualitative analysis require some quantitative substantiation to ascertain, for example, whether there are significant differences between the accuracies of the filtered datasets. The results from the quantitative analysis are presented in Section 8.4.1 along with the results from the other two filtering methods.

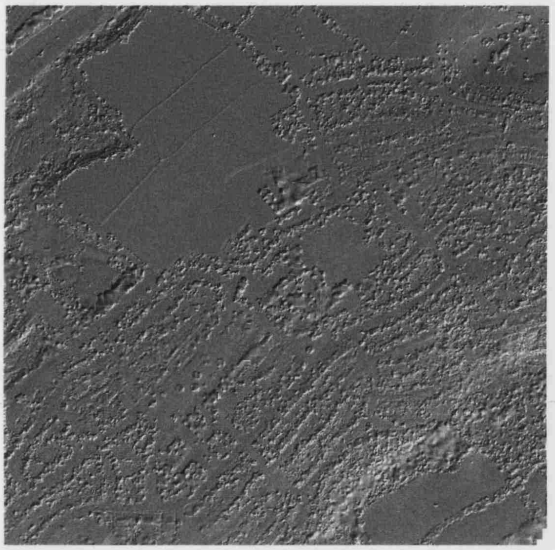

Figure 8.36: DTM created using bicubic gridded input data at 0.5m resolution with the default IRI method

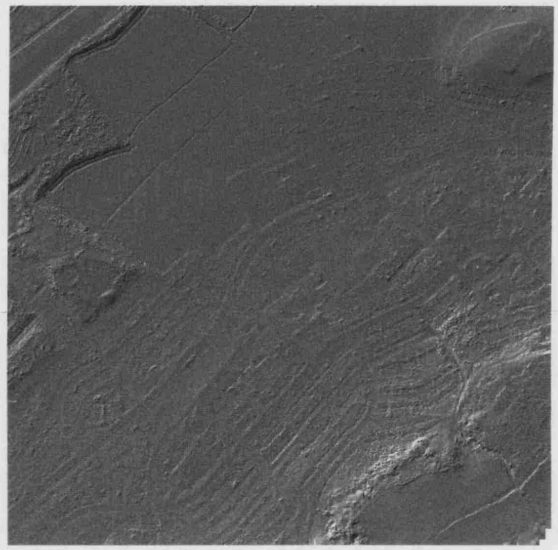

Figure 8.37: DTM created using nearest neighbour gridded input data at 0.5m resolution with the default IRI method

8.3.3 Adaptive TIN Models (ATM)

The general principles of the Adaptive TIN models algorithm was introduced in Chapter 3. It was chosen for comparison in this study on the basis of the favourable results produced in Sithole and Vosselman (2003) and the fact that the algorithm is commercially available within the TerraModeller software suite. The method uses the discrete points as inputs to the filtering process, which are then connected in a TIN during the identification of candidate terrain points. An initial sparse TIN is identified and iteratively densified in accordance with parameters defined by the user. Whilst the overall methodology for this algorithm was created by Axelsson (2000), the specific parameters for its implementation must be defined by the user. The parameters used in this study are described below, again two different settings were used, so that the sensitivity of the algorithm to these changes could be assessed.

The first stage of the process involves the identification of any apparently erroneous low points. These were identified on a point by point basis, whereby each point was compared with every other within a 5m radius. Any point lying more than 0.5m below any other point in a 5m radius was considered to be erroneous, and

was classified as a low point. However, classifying in this way, on a point by point basis, does not account for the occurrence of groups of erroneous points. In order to eliminate any such groups of points a second routine was performed to search for groups of points which were 0.5 meters lower than other points in the vicinity. Erroneous low points may occur in laser scanning data as a result of miscalibration of the system: however, this is usually corrected before the data are supplied to the user. In addition, erroneous low height values are often found over water areas, or highly reflective surfaces such as windows and metal. In such areas no reflectance value at all may be recorded: some vendors record such instances as a 'no data' values by assigning them a very low value. For the Bristol test area, there proved to be no such low points, probably because these had either been filtered by processing of the data prior to supply, or because no erroneous low points were created in the data capture process.

In order to establish the initial ground model, potential terrain candidates were chosen from the remaining points. In order to choose these, a maximum building size was defined as being 150m^2. A search window the size of the largest building was then defined and it was assumed that the lowest point within any search window must be terrain. The resulting points were connected together in a TIN to represent the initial ground surface, and this was used to ascertain whether remaining points represented part of the terrain. Points were added to the terrain if they met certain requirements, which were defined in terms of threshold parameters. In this way the TIN was progressively densified upwards. The process was iterated until all points were classified as ground or object.

The chosen iteration parameters determined how close a candidate terrain point had to be to a triangular plane, in order for it to be accepted as part of the TIN. Values for these parameters were estimated using all statistics derived from all data points. Parameters included the estimation of maximum building size (150m^2), terrain angle, iteration angle and distance. The terrain angle was estimated using a slope map of the area. Values for the average angle ranged between 5 and 30 degrees.

The second parameter, iteration angle, is defined as the maximum angle between any point, its projection triangle plane and the closest triangle vertex (see Figure 8.38). Small values of the iteration angle mean that the threshold for including terrain points is very low, in other words the routine is less likely to follow changes in the point cloud, that might represent small undulations in terrain or hits on low vegetation (TerraSolid Oy 2004). Values for the iteration angle ranged between 3 and 4 degrees, as the terrain in the study region was relatively flat (although not uniformly so). The third parameter to be defined is that of iteration distance which is the distance between the candidate point and the surface normal (see Figure 8.38). Typical values for iteration distance for the study region ranged between 1.2 and 1.5m.

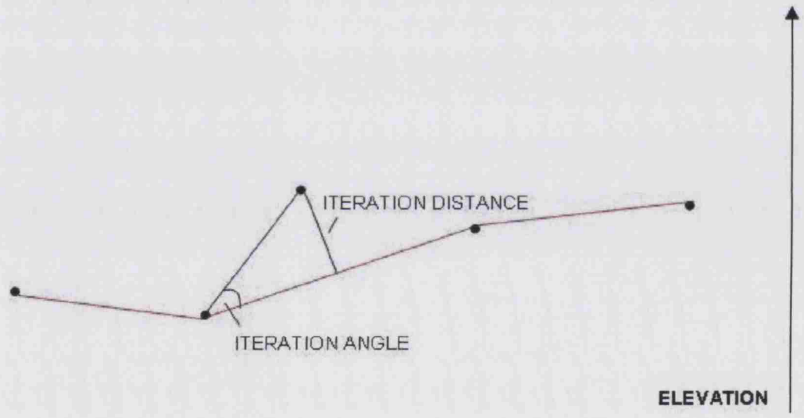

Figure 8.38: The role of the iteration angle and distance in the ATM algorithm

In order to choose the parameter settings appropriate for this investigation a pilot study was conducted in which DTMs were created using different parameter settings. The resultant DTMs were visually inspected and the success of the algorithm asessed qualitatively in order to ascertain which of the parameter settings appeared

to be the most suitable. The DTMs created in this pilot study are presented in Appendix H. The results of the pilot study investigation revealed the following: in terms of choosing the optimum terrain angle value, an angle of 5^o was deemed to be insufficient as it led the inclusion of many above ground objects as terrain, whilst an angle of 20^o was too great and led to a scouring of the landscape during filtering. Based on these observations a terrain angle of 15^o was chosen, given that the DTMs created using this setting appeared to suffer from the least amount of scouring. The other parameter settings were chosen in a similar qualitative way. To assess the sensitivity of the ATM algorithm the results from the filtering using the modified parameter settings were compared with those produced using the software default settings. Table 8.2 details the parameter settings used for both the default and the adapted methods.

Table 8.2: Parameter Settings for default and adapted ATM algorithm

Parameters	Default Settings	Adapted Settings
Maximum buildings size (m^2)	60	150
Terrain angle (degrees)	75	15
Iteration angle (degrees)	6	4
Iteration distance (m)	1.2	1.2

Images of the resultant datasets for one of the parameter settings are shown in Figures 8.39 and 8.40. The black regions of the terrain models were those areas where there were insufficient ground points during the computation of the TIN given that there was a constraint on the maximum size of the triangle permitted. In the images the boundary between the coloured and the black areas is not necessarily representative of any boundary between the ground and an object: it is simply the edge of a triangle. It was noted that in Figure 8.39 and in Appendix H some of the resultant filtered terrain point datasets exhibited an edge effect in the lower right segment. It was observed that this only occurred where the terrain angle was set at 5 degrees, at higher terrain angles no such feature was noted. The terrain angle is the steepest allowed slope in the ground terrain. As such, the edge effect was caused by the steeper terrain in these areas. For those settings which used the lower terrain

angle of 5 degrees the ground classification in these steepper areas was unsuccessful. The edge effect reflects the underlying change in slope. This edge effect did not effect the main investigation, which used the terrain angle of 15 degrees for the filtering.

Figure 8.39: Terrain/ground points only. Coloured by elevation (m)

Figure 8.40: Resultant DTM created from ground points. Key in (m)

There are, as with any method, a number of assumptions made in the ATM algorithm. Firstly, it is assumed that there can be one global value for the terrain

angle within the study region. Given the variation in terrain in this investigation, this was found to be problematic for filtering some of the datasets. Secondly, it is assumed that passing a kernel the size of the maximum building is optimal, while it may in fact be better if a smaller kernel were to be used. Whilst this may introduce some error in filtering in the region of the maximum building, it was thought that using a smaller kernel may permit the inclusion of more candidate terrain points in the rest of the study region and so improve overall classification accuracy. This characteristic of the algorithm was particularly relevant for the study region used in this investigation, in which the largest building was much greater than the average building size.

A small representative selection of images of the resultant filtered datasets are shown in Figures 8.41 to 8.46. There was much less obvious difference between the forms of the DTMs than had been noted for the Geometric or IRI algorithms. In particular, it was difficult to discern any differences between the DTMs created using the various interpolation methods, an observation which contrasts with the results for the IRI filtering. The lack of difference could either have been caused by the display environment of the software used for the ATM algorithm (which was different from the previous methods), or may have been caused by the algorithm itself if some 'scouring' had occurred. Scouring occurs where both objects and ground may have been removed in the filtering process: whilst this produces a much smoother DTM it also introduces both classification and height accuracy errors. One noticeable difference for the ATM was that the bicubic surface appeared more 'pitted' than the others.

The appearance of the DTM produced from filtering of the raw data (Figure 8.41) suggested that it had been slightly more successful than the DTMs created from interpolated data (Figures 8.42 to 8.44). In general the ATM DTMs appeared to contain less surface detail than those produced by the IRI method, given the lack of building footprints and remaining above ground objects: this was also likely to be a consequence of scouring during filtering. The analysis of the proportion of Type

I and Type II errors for all three of the filtering methods (see Section 8.4.3) should help ascertain if this is the case. Finally, there also seemed to be very little difference between the forms of the surfaces created from input data of different resolutions, as shown in Figures 8.45 and 8.46. These qualitative observations regarding the form of the filtered surfaces require substantiation by quantitative analysis. This analysis is described in Sections 8.4 and 8.4.1 below.

Figure 8.41: DTM created from filtering of raw data using default ATM settings

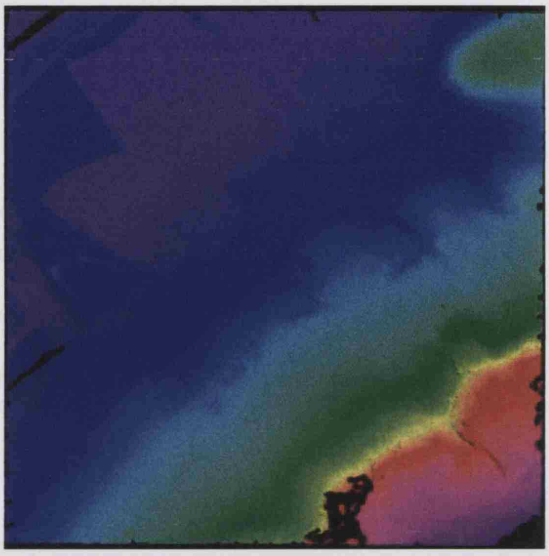

Figure 8.42: DTM created from Bilinear data at 1m using default ATM method

Figure 8.43: DTM created from Bicubic interpolated data at 1m using default ATM method

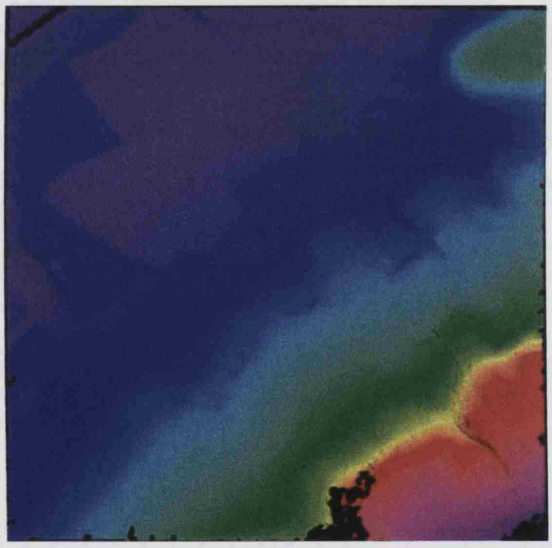

Figure 8.44: DTM created from Nearest Neighbour data at 1m using default ATM method

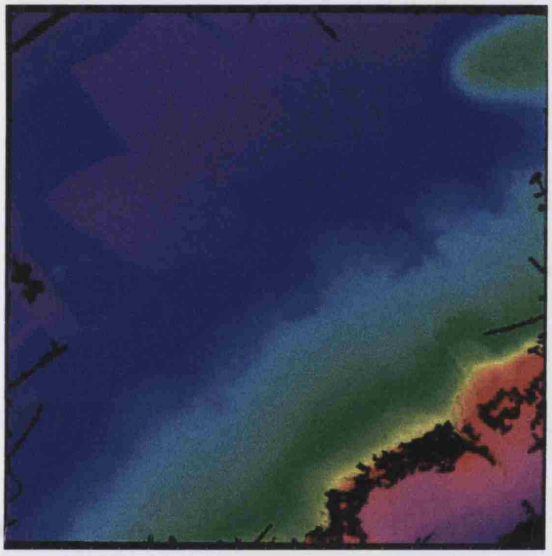

Figure 8.45: DTM created from Bilinear data at 0.5m using default ATM method

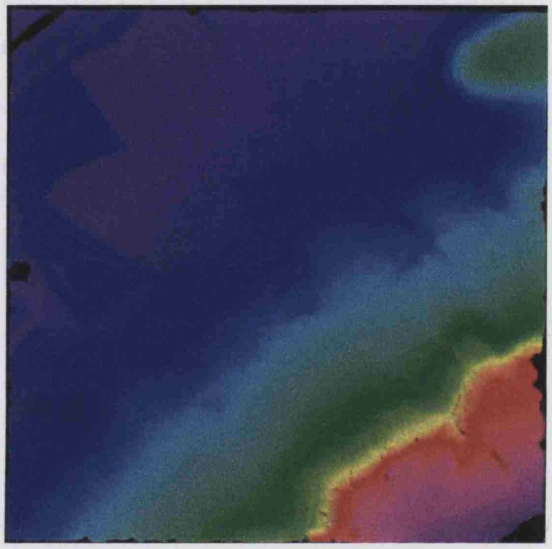

Figure 8.46: DTM created from Bilinear data at 2m using default ATM method

8.4 Quantitative Comparison of the Filtered Datasets: Classification Accuracy

Quantitative assessment of the various filtered datasets should address two components of filtering. These are the classification accuracy and the accuracy of the modelled terrain heights. In this section the classification accuracy of each of the filtered datasets is assessed.

In order to assess the accuracy of the filtering classification of ground points, a sub-sample of each of the datasets was taken and compared visually with an aerial photograph (see the example in Figure 8.47). A visual comparison was used in place of an automatic classification of the image, as any such classification would accommodate some degree of error which would then be transferred to and affect the validation results (as was found in the investigation presented in Chapter 6). At a 95% confidence level with a confidence interval of 10 the required sub-sample size was 96 points. These points were randomly selected from each of the filtered datasets. The sample sites were chosen for each of the datasets using this random approach, and as a result each of the samples contained different locations. In this respect the accuracy classifications were independent.

In order to assess the success of the classification, contingency tables were calculated for each dataset. The contingency tables were then compared to assess whether the differences between each were statistically significantly different. Many previous studies which have investigated the statistical significance of differences in classification accuracy have based their analysis upon the comparison of accuracy expressed in terms of the kappa coefficient of agreement. However there are a number of concerns which have been presented regarding the use of this measure (e.g. Foody (2004); Stehman (1997)). As discussed in Chapter 4, given the contentiousness of using the kappa coefficient it may be preferable to use alternative metrics for accuracy assessment, such as those presented in Foody (2004), one of which is the z-test for independent samples.

Figure 8.47: An orthorectified aerial photograph of the study area, with one of the filtered point datasets overlain for purposes of comparison of classification. The point dataset shown (pink crosshairs) is that created from the SCOP DTM method for the bicubic 1m data

The z-test is a test for the significance of the difference between two proportions, and is estimated from

$$z = \frac{\frac{x_1}{n_1} - \frac{x_2}{n_2}}{\sqrt{p(1-p)(\frac{1}{n_1} + \frac{1}{n_2})}} \tag{8.1}$$

where x_1 and y_1 represent the number of correctly allocated cases in two independent samples of size n_1 and n_2 respectively, and $p = (x_1 + x_2)/(n_1 + n_2)$.

The statistical difference in accuracy between the two classifications was evaluated through a comparison of the calculated value of z (calc-z) with the expected value (tab-z).

A worked example of the use of the z-test for assessing differences in classification accuracy is shown below. In the example the z-test is used to ascertain whether there is a difference between the filtering classification accuracies of the bicubic and nearest neighbour datasets. Whilst it was shown in Chapter 6 that the bicubic introduced less error than the nearest neighbour algorithm during restructuring, it remains a possibility that the smoothing of the surface introduced by the bicubic method may in fact be detrimental in subsequent analysis and so create a difference in filtering results. In other words, the null hypothesis for this investigation states that there is no difference in the filtering accuracies of the bicubic and nearest neighbour datasets, conversely the alternative hypothesis states that there is a significant difference in accuracy between the two datasets.

The contingency tables showing the classification accuracies of ground points in the bicubic and nearest neighbour surfaces using the Iterative Robust Interpolation (IRI) method are shown in Tables 8.3 and 8.4.

A comparison of the contingency tables using the z-test (Equation 6.1) yields the

following

$$z = \frac{131/189 - 134/160}{\sqrt{0.76(1-0.76)(1/189+1/160)}} \tag{8.2}$$

$$z = \frac{-0.14}{\sqrt{0.05)}}$$

$$z = -3.14$$

$$|z| = 3.14$$

Where p is derived as:

$$p = \frac{(x_1 + x_2)}{(n_1 + n_2)} \tag{8.3}$$

$$p = \frac{(x_1 + x_2)}{(n_1 + n_2)}$$

$$p = \frac{(131 + 134)}{(189 + 160)}$$

$$p = 0.76$$

As the calculated value of z (calc-z) (3.14) is greater than the tabulated value of z (tab-z) (1.96) the null hypothesis should be rejected, and the alternative hypothesis

Table 8.3: Contingency Table for the adapted IRI filtering method for the Bicubic 1m dataset

		Truth Data Ground	Above Ground	**Totals**
Classification Data	Ground	79	6	85
	Above Ground	20	55	75
	Totals	99	61	160

be accepted at the 5% significance level for a two-tailed test, indicating that there is a significant difference between the accuracy of the classifications in bicubic and nearest neighbour surfaces for the IRI method.

This result was expected following initial qualitative analysis of the DTMs created from the filtered datasets. The original surface is shown in Figure 8.48. The filtered surfaces for the bicubic and nearest neighbour are shown in Figures 8.49 and 8.50 below. It can be seen below that the cubic surface in Figure 8.49 contains far more obviously erroneous features than the nearest neighbour surface (Figure 8.50) in which most of the above ground features have been successfully removed. This observation is corroborated by the z-test results.

It was necessary to make a number of assumptions in the classification validation experiment. Firstly, it was assumed that there had been no change in the surface form of the landscape between capture of the laser data and the aerial photograph. In a similar way it was assumed that there were no significant differences in the accuracy of the two datasets, and that both had been correctly geo-referenced. In other words, any differences between the two were purely a product of the algorithms rather than any differences between the datasets. Finally, it was assumed that the manual classifications of the aerial photograph, which were used to assess the accuracy of the filtered datasets, were themselves reliable. Following the definitions of research reliability offered in Chapter 3, it was considered that these classifications could be deemed reliable if they were repeatable. For this reason 198 of the points classified within this investigation were classified again by four research assistants working independently. The classifications for each were compared with the original

Table 8.4: Contingency Table for adapted IRI method for the Nearest Neighbour 1m dataset

		Truth Data Ground	Above Ground	**Totals**
Classification Data	Ground	89	7	96
	Above Ground	51	42	93
	Totals	140	49	189

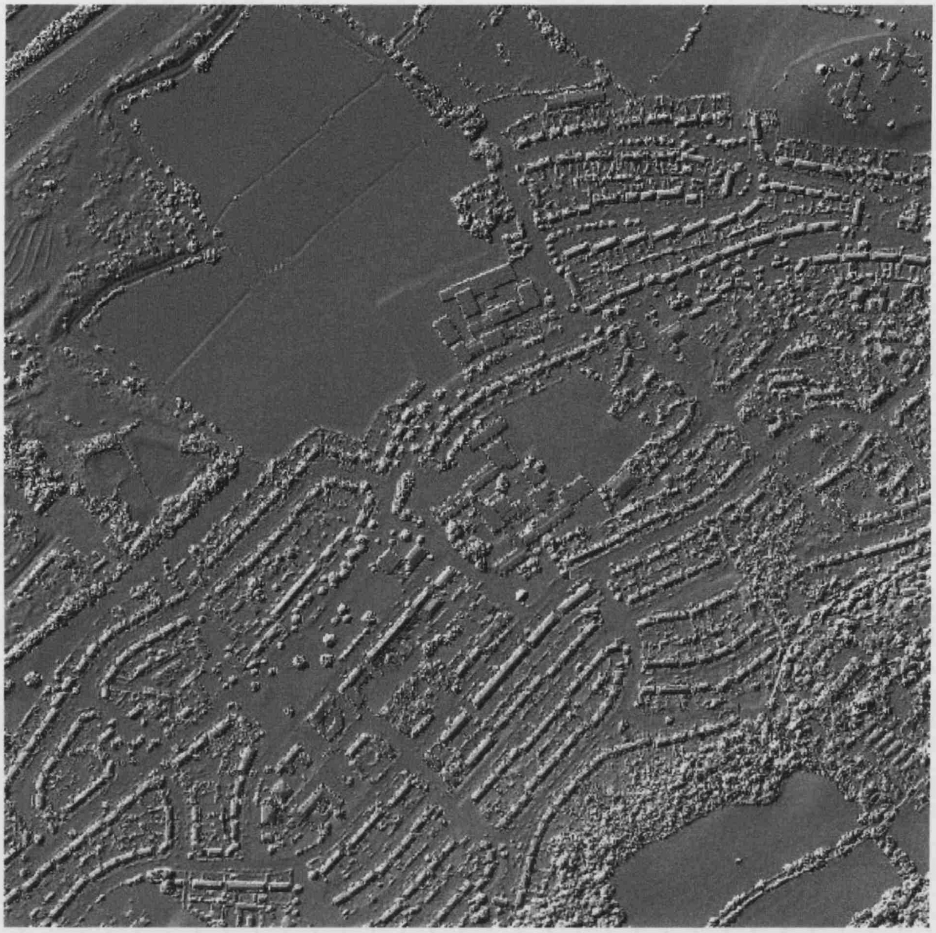

Figure 8.48: Original surface (DSM) before filtering of above ground points

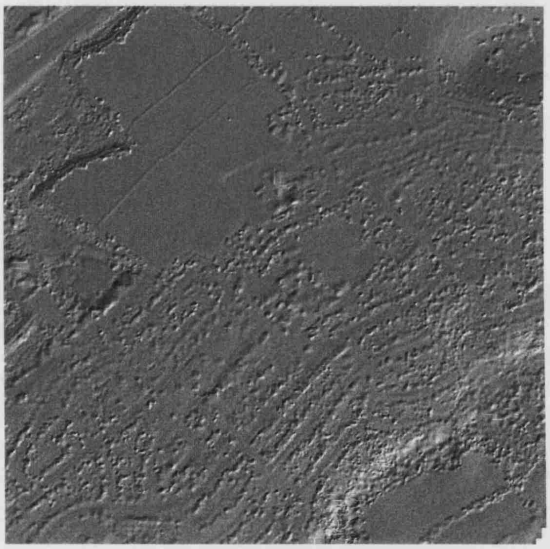

Figure 8.49: DTM created from filtered bicubic 1m surface

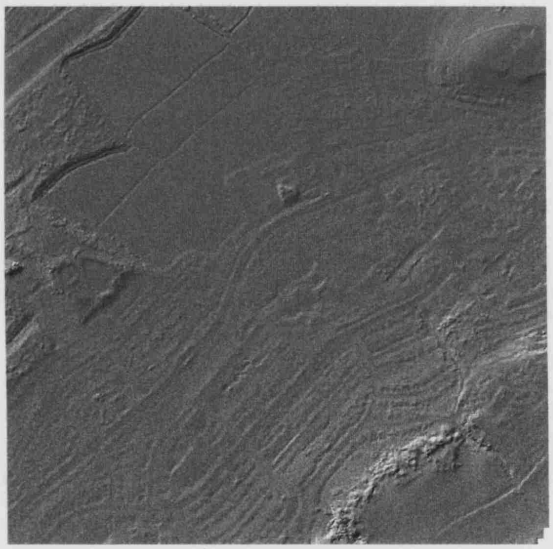

Figure 8.50: DTM created from filtered nearest neighbour 1m surface

classification to assess reliability. The full results are presented in Appendix I, and Table 8.5 summarises the correspondence between them.

Table 8.5: Results for Reliability Test for Classification Accuracy Assessment

Research Assistants	Percentage Similarity With Original Classification
1	80.73
2	68.23
3	71.35
4	71.88

The datasets compared in this reliability investigation were related, as they were all from the same 198 sites. Given this a McNemar (Agresti (1996)) test was used in line with Foody (2004). The McNemar test states that

$$z = \frac{f_{12} - f_{21}}{\sqrt{f_{12} + f_{21}}} \tag{8.4}$$

Where f_ij indicates the frequency of sites lying in confusion matrix element i, j (after Foody 2004).

Foody (2004) notes that for some remote sensing applications an adaptation of this statistic, which is based on a chi-square (χ^2)) distribution, is more common. However he also observes that this adaptation is not suitable for a two-tailed test, and as such was deemed to be unsuitable for this investigation.

The results from the McNemar test indicated that there was no significant difference between the research assistant classifications and the original classifications at the 5% significance level (see Appendix I for calculations). It was concluded, therefore, that the classifications were repeatable indicating that the technique of manually comparing the filtering classifications with the aerial photograph was reliable.

8.4.1 Results and Discussion: Overall classification accuracy

In this section all of the results of the investigations assessing the accuracy of the filtering classifications are presented. This accuracy is assessed both in terms of overall and of producer's accuracy. The analysis of producer accuracy specifically shows the success of the 'bare earth' classification, which is of most relevance to the filtering stage of post-processing.

The overall accuracies of the datasets are shown in Tables 8.6 to 8.8. The most successful algorithm appeared to be the adapted geometric method which produced classification accuracies ranging between 84% and 93.7%, whilst the poorest performer was the IRI adapted method which produced accuracies as low as 49.7%. In general, classification accuracies using the raw data were in fact lower than those produced by using some forms of interpolated data. In particular data interpolated using the nearest neighbour method were found to produce the highest classification accuracies for many of the algorithms (including both forms of the ATM and geometric methods).

The reason for this, counter-intuitive, finding was that the interpolation procedures, especially the nearest neighbour technique, produced artefacts in the surface such as step edges. These do not occur in the raw dataset. Most filtering algorithms look for rapid changes in slope, such as these step edges, as indicators of surficial features which are to be removed from the dataset. Thus, ironic as it may seem, these interpolation artefacts will actually make the filtering algorithm more efficient. In other words, by introducing errors during interpolation, the overall filtering errors may be reduced. The usual view of errors in geographical datasets is that they are bad. This research suggests that this is not always the case.

Table 8.6: Overall accuracy results table for IRI method

Algorithm	Dataset	Grid Space	Overall Accuracy
IRI default	**Cubic**	2m	69.31%
		1m	73.56%
		0.5m	83.31%
	Linear	2m	92.35%
		1m	96.00%
		0.5m	89.17%
	Nearest Neighbour	2m	83.75%
		1m	83.85%
		0.5m	79.63%
	Raw data	Original point density	81.53%
IRI adapted	**Cubic**	2m	54.90%
		1m	52.38%
		0.5m	49.26%
	Linear	2m	88.11%
		1m	80.77%
		0.5m	78.57%
	Nearest Neighbour	2m	80.95%
		1m	73.79%
		0.5m	69.18%
	Raw data	Original point density	67.57%

Table 8.7: Overall accuracy results table for ATM method

Algorithm	Dataset	Grid Space	Overall Accuracy
ATM default	**Cubic**	2m	78.93%
		1m	85.31%
		0.5m	81.01%
	Linear	2m	83.60%
		1m	84.19%
		0.5m	84.82%
	Nearest Neighbour	2m	84.50%
		1m	86.72%
		0.5m	82.42%
	Raw data	Original point density	79.70%
ATM adapted	**Cubic**	2m	81.62%
		1m	72.56%
		0.5m	59.11%
	Linear	2m	80.08%
		1m	85.71%
		0.5m	83.14%
	Nearest Neighbour	2m	86.36%
		1m	85.93%
		0.5m	83.03%
	Raw data	Original point density	83.21%

Table 8.8: Overall accuracy results table for Geometric method

Algorithm	Dataset	Grid Space	Overall Accuracy
Geometric default	**Cubic**	2m	79.08%
		1m	80.21%
		0.5m	76.00%
	Linear	2m	76.14%
		1m	76.40%
		0.5m	77.74%
	Nearest Neighbour	2m	79.39%
		1m	87.96%
		0.5m	82.35%
Geometric adapted	**Cubic**	2m	86.47%
		1m	89.59%
		0.5m	86.08%
	Linear	2m	88.06%
		1m	88.68%
		0.5m	84.00%
	Nearest Neighbour	2m	92.73%
		1m	92.31%
		0.5m	93.48%

Specific observations for each of the three algorithms are noted below.

8.4.1.1 The IRI Method

For the IRI method the highest classification accuracies were produced using the bilinearly interpolated data, and the lowest accuracies by the bicubic method. This was surprising given the results of the investigation presented in Chapter 6 which indicated that the bicubic method produced some of the lowest amounts of error of all the interpolation methods for data restructuring. This apparent anomaly was considered to be a product of the workings of the IRI method in the early stages of the algorithm. Recalling from Section 8.3.2 (page 239), in the first stage of the IRI algorithm an approximate surface is computed using some of the input points. This approximation is done using linear prediction, which, given the interpolation method and the fact that only a few points are used, will produce a very smooth surface that is likely to be reasonably similar to the bicubic surface. The input points are then all compared to this approximate surface and each point given a weighting according to its distance to the surface. The weights are used to compute the next surface, where points with high weights (i.e. low residuals) attract the surface whilst points with low weights will have little influence on the form of the surface and as such will be excluded from the ground point class. Clearly the differences between the input points and the approximated surface will be very different for each of the datasets: the nearest neighbour points for example will be very different from the smoothed approximate surface, and so will produce large residuals, and therefore lower weights. Conversely, the bicubic input points may be very similar to the approximate surface, and so produce very small residuals (and therefore a high weighting value) meaning that most points, including those representing above ground objects, will be included within the ground point classification. In other words, because of the similarity of the bicubic points to the approximate surface the algorithm has difficulty distinguishing between ground level and above-ground points. Whilst this difference between the bicubic and the nearest neighbour points

is reasonably straight-forward it is, perhaps, not so apparent why the bilinear input points should produce the highest classification accuracies given that the interpolation method used in the IRI algorithm is a linear predictor. It is considered here that this is related to the reduced number of points which are used by the IRI algorithm, which when used by the linear predictor produce a surface which is much smoother than that created by using a bilinear interpolator from the raw data. It is proposed here that this overly smoothed approximate surface, is more like the bicubic points than the bilinear.

The results in Table 8.6 also show that the default method for the IRI algorithm produced higher overall accuracy classifications than the adapted method, which was slightly surprising given that on visual inspection the terrain models produced by the adapted method appeared more successful - indicating the importance of a quantitative accuracy assessment (as advocated by Foody 2004).

8.4.1.2 The ATM Method

In terms of the classification accuracies for the ATM algorithm, it can be seen that there is little difference between the accuracies of any of the datasets. This suggests that data structure influences the accuracy of this algorithm to a far lesser degree than for the IRI method. The modified ATM algorithm produced slightly better results than the default method, which was not surprising given that the modified algorithm was tailored for use with the study area dataset. The significance of this difference is tested for later in this chapter.

8.4.1.3 The Geometric Method

The geometric method also exhibited a high success rate for classification accuracies, with a low range of fluctuation. In particular, the adapted settings for this algorithm produced consistently good results for all input datasets.

On the basis of these observations it is clear that a user wishing to reduce the amount of overall classification error in a DTM should interpolate the raw data using the nearest neighbour method and then use the geometric algorithm, with adapted settings, to filter the gridded data. These results are based on the data used for this study, which whilst variable in relief and landcover cannot be taken to represent all terrain. As such further work should investigate the transferability of these observations to other, less common, terrain and land cover types.

8.4.2 Results and Discussion: Producer's Accuracy

The overall accuracy results provide a measure of the general success of the filtering algorithm. However just because an algorithm has a high classification accuracy this does not mean that every category was successfully classified at the same rate. For an indication of the accuracy of each of the classifications, the producer's accuracy was calculated by dividing the number of correct classifications for each category by the number of classifications in the category. In terms of the contingency tables used in this investigation the producer's accuracy is the number of correctly classified points/pixels divided by the column total. The results of these calculations are shown in Figure 8.51. The full tables of results are presented in Appendix J.

The bar charts show that for almost all of the algorithms, the nearest neighbour dataset produces the highest accuracy for ground classification, and frequently this accuracy is higher than that of the filtered raw dataset. Conversely, the cubic dataset consistently produces some of the least successful results. This is a very important finding, as it refutes the long held assumption that restructured data will always produce lower accuracy datasets. In the case of the nearest neighbour datasets this is clearly not the case.

In terms of differences between the algorithms, the geometric algorithm was found to produce some of the highest accuracies, with many greater than 90%. This result was surprising given the relative simplicity of the geometric algorithm, and the fact

that it is a grid based method. Traditionally these have been thought to produce lower accuracy datasets than the point based methods.

In summary there are two important findings from this investigation. Firstly that restructured data can produce higher classification accuracies than could be achieved using the raw data. Second, the results obtained using the grid based filtering methods can produce higher accuracies than point based methods.

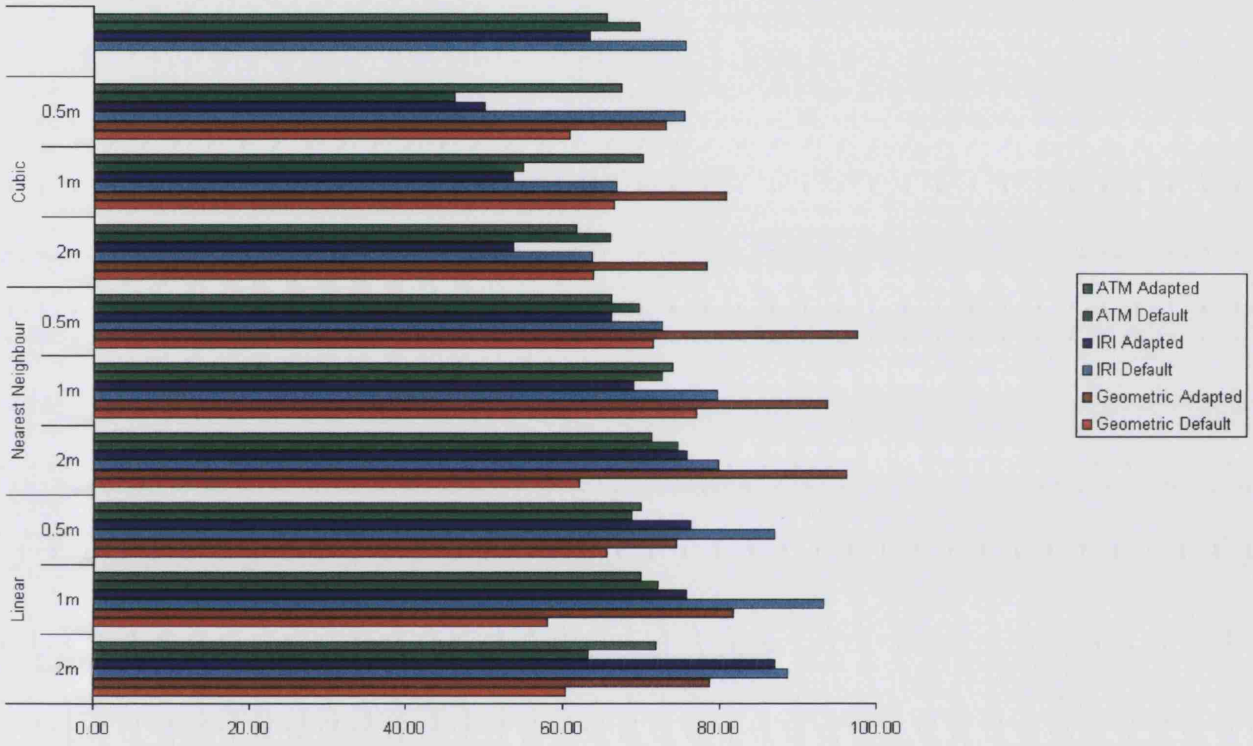

Figure 8.51: Bar chart showing the calculated producer's acuracies for all three filtering methods

8.4.3 Type I vs Type II Classification Errors

All filtering algorithms make a distinction between above ground and bare earth based on the assumption that certain characteristics are associated with the former and some with the latter. This assumption while often valid does sometimes fail (Sithole and Vosselman 2003) as evidenced by the results in the previous section. As a result there is a trade-off between making Type I (rejecting bare earth points) and Type II errors (accepting object points as ground). It is likely that different input datasets, parameter settings and filtering methods will all produce different patterns of both types of error. The spatial patterns of filtering errors are explored qualitatively here in order to give some further indication of the characteristics of the different algorithms.

Figures 8.52 to 8.55 show the spatial patterning of errors for three example datasets, which are the classifications for the cubic 1m data created by the three filtering algorithms. The Figures show the location of both Type I and Type II errors. The patterns of error for each class reveal information about the strengths and weaknesses of the algorithms which are not apparent from the statistical investigation presented in the previous section.

In the ATM dataset (Figure 8.52) there were substantially more Type I than Type II errors. In other words, many 'true' ground points were being mis-classified as 'above ground' objects. These points would be eliminated from the dataset during filtering, and in the computation of the resultant DTM the ground surface would be much lower in the model than it should be. This effect is termed 'scouring'.

There was no obvious pattern in the location of the above ground errors, although some of these errors did appear to occur in close proximity to some above ground objects. Of more concern were the errors which occurred amidst clear regions of bare earth in the study area (such as those in the main field). Whilst it is possible that these points within the original laser dataset did in fact occur above ground

Figure 8.52: Aerial photograph of study area overlain with classification errors. Red crosshairs (Type II) are errors in ground class, blue crosshairs (Type I) show errors in above ground classification for the ATM algorithm. Note the predominance of Type I errors

(such as may have happened if the laser beam had coincided with a bird or raindrop and therefore recorded an above ground height at this location), this explanation is extremely unlikely to have accounted for all of these errors, and it is more likely that they have been caused by some problem within the algorithm. The Type II errors appeared to mostly occur over vegetation. However, there is much less certainty regarding whether or not these are in fact errors at all. In vegetated areas the laser may have penetrated through to the ground rather than hit the vegetation itself. However, in terms of classification accuracy it was assumed that any laser point occurring over the vegetation class would have recorded the vegetation height and not the bare earth height.

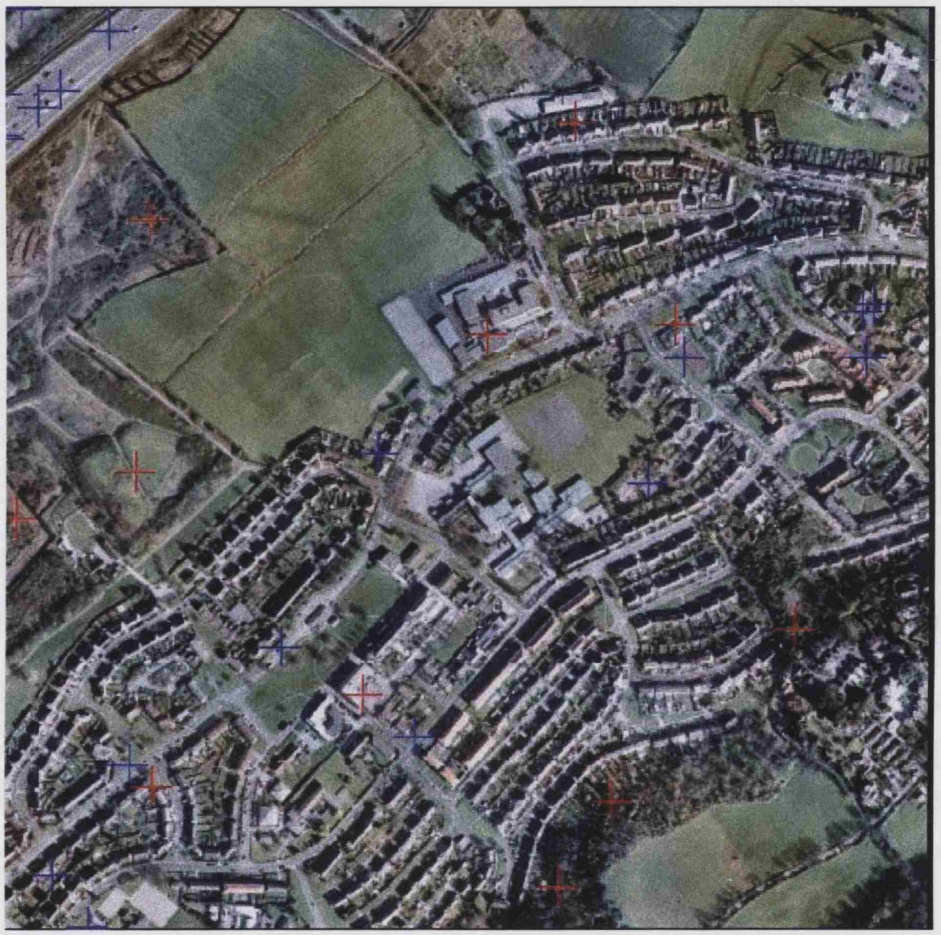

Figure 8.53: Aerial photograph of study area overlain with classification errors. Red crosshairs (Type II) are errors in ground class, blue crosshairs (Type I) show errors in above ground classification for the geometric algorithm

Figure 8.54: Aerial photograph of study area overlain with classification errors. Red crosshairs (Type II) are errors in ground class, blue crosshairs (Type I) show errors in above ground classification for the IRI algorithm

Figure 8.55: Showing the error in the bare earth classification on the multi-roofed building. Note the mis-classified point (red crosshair) in the centre-left of the image. This point is clearly positioned over of part of a building but has been erroneously classified as bare earth by the algorithm

The pattern of errors for the geometric algorithm exhibited very different characteristics. The occurrence of Type I and Type II errors are shown in Figure 8.53. It is clear that in terms of the overall number of errors there are far fewer than noted in the ATM dataset, with both types of errors appearing to occur over an object in the associated aerial photograph. The Type II errors all occurred close to an above ground object suggesting that the filtering algorithm had encountered some problems determining the boundaries of 'above ground' objects.

In some instances misclassifications had occurred over complex objects such as those shown in Figure 8.55 which shows a building roof with many levels. In Figure 8.55 the incorrect bare earth point clearly lies on the lower roof of the building rather than on the bare earth. This error occurred as a consequence of the assumptions of the geometric algorithm, in which it is assumed that large differences in heights between neighbouring points/pixels indicate a change between object and bare earth. For this reason the lower point has been classified, incorrectly, as bare earth. This

problem with the algorithm could be amended if RGB values for each laser point could be used in addition to height values, and so reduce the possibility of this happening in future.

In terms of the pattern of the above ground errors, again it appeared that these were more clustered than those in the ATM dataset. In the geometric results some clear patterns were apparent, such as those errors occurring close to groups of buildings where clearly the errors were the result of uncertainty regarding the extent of the buildings/trees. There was also a cluster of errors in the top left of the dataset over the motorway. It was thought likely that these were not errors caused by the algorithm, but were instead the product of difference between the laser dataset and the aerial photograph. At the time of laser data capture there may have been (large) vehicles on the road, which the algorithm classified as above ground objects, whilst in the aerial photograph these temporary features were absent thus creating a difference in the classifications. These errors should be treated as 'fuzzy-errors' in the sense that we cannot be certain that they actually are errors within the algorithm, and as such should potentially be eliminated from the computation of statistics.

Finally, the errors from the IRI cubic 1m dataset are shown in Figure 8.54. Again, the patterns of errors reveal certain characteristics about the IRI algorithm which are different from the characteristics of the other algorithms. In particular, there seemed to be no dominance of either Type I or Type II errors, with both occurring rather indiscriminately over the study area. The above ground errors (Type II) did not appear to display any particular pattern, and many occurred on open ground in which the algorithm should not have experienced difficulty during classification. The reason for the occurrence of errors in these areas is somewhat enigmatic. The only pattern discernable in the distribution of bare earth (Type I) errors was that most appeared to have occurred over vegetation: again these were considered to be 'fuzzy errors'.

As can be seen from the figures, all three algorithms displayed a tendency to minimise Type I errors, in other words the filters attempt to model the bare earth as realistically as possible, with the trade-off that many object points may remain in the dataset. This characteristic of the ATM algorithm has previously been noted by Sithole and Vosselman (2004). This is an interesting observation, given that previous studies have found that other filtering algorithms have concentrated more on the elimination of Type II errors. The clear implication is that there is no overall consensus amongst the authors of the algorithms as to which type of error is the more costly. The bar charts (Figures 8.56 to 8.58) show the percentage error for each of the datasets shown above. What is clear is that the geometric algorithm consistently produces the lowest amount of error, and is the most successful at minimising Type II errors, whilst the ATM method is the most successful at reducing Type I errors.

The findings yielded by examining the spatial pattern of both types of errors are therefore important, as they reveal more information about the characteristics of the algorithms than the statistics alone can show. Examples include the clustering of Type II errors over the motorway for the geometric algorithm. As stated above, these are probably not errors at all, and if this is assumed then the success of the filtering in this algorithm is increased. Similarly the coincidence of bare earth errors in areas of vegetation also influences the perception of algorithm success, given that classifications of laser points in these regions are uncertain the user should not place too great a confidence in the occurrence of such errors. In other words, an algorithm should not be rejected on the basis of statistics alone, as the spatial location, patterning, and the spread across categories of the errors may significantly influence the decision. In this respect, the map of errors may be at least as important as the computation of (global) accuracy statistics for the user of laser data. Methods for visualising this error and uncertainty are discussed in Chapter 10.

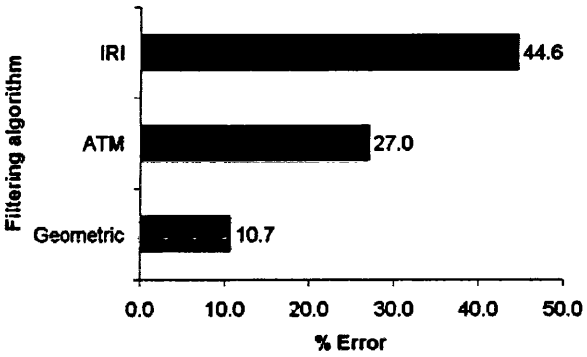

Figure 8.56: Percentages of total error

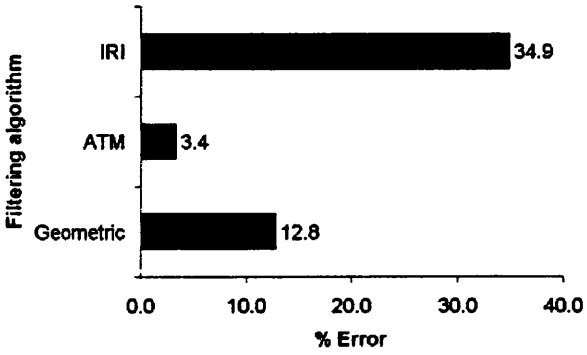

Figure 8.57: Percentages of Type I error

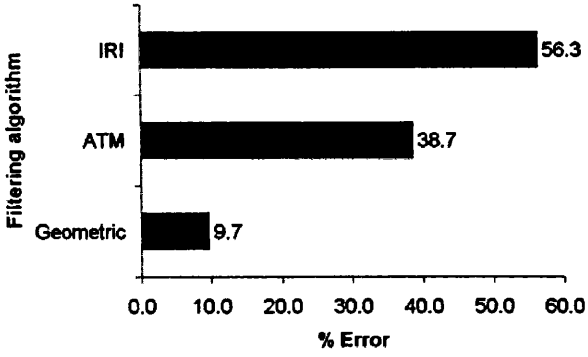

Figure 8.58: Percentages of Type II error

8.5 Quantitative Comparison of Filtered Datasets: Accuracy of the Height Modelling

The second component of the filtering accuracy which required quantitative assessment was the modelling of the ground heights. For this it was necessary to use reference datasets with which the filtering results could be compared.

The filtered datasets were first compared with the photogrammetrically derived reference DTM (Figure 8.60). In order to establish differences each of the test datasets was subtracted from the reference model. The resulting images showed the areas of greatest variation between the filtered datasets and the reference model. Whilst the outputs from both the IRI and the geometric methods were grids and therefore straight forward to analyse as rasters, the output from the ATM method was datasets of discrete points which could not be used for differencing in this way. In order to use a standard methodology for all datasets, and so increase the validity of the investigation, the discrete ATM points were exported as a grid at the required resolution for each outputted dataset. It is likely that this process may have introduced some error into the dataset. The differences were analysed qualitatively, and the results are documented in Section 8.5.1.

In order to analyse the pattern of errors within the filtered height models, a relatively simple method was adopted (a justification for which was offered in Chapter 4). The methods used are based on qualitative descriptions of the model differences and spatial patterns of errors, which are combined with quantitative statistical analysis in order to determine whether, and in what ways, the datasets are significantly different. In effect this method is an extension of those described by Li (1992), Fischer and Dowman (2002), and Dowman and Fischer (2003).

The values in the filtered surfaces were also compared with each of the RTK GPS readings at the specified ground control points. The differences between the GPS height values and the filtered datasets were calculated, as were the descriptive sta-

tistics for these residuals. These results are presented in Section 8.5.1.

8.5.1 Results and Discussion: Image Differencing

The image differencing revealed some substantial variations between the filtered datasets (an example of which is shown in Figure 8.59) and the reference model (Figure 8.60). Some of the resultant images are shown in Figures 8.61 to 8.63. In this analysis it was important to consider both the difference images and the descriptive statistics for each dataset, given that each of the image differences used different colour scales which were representative of the spread of values within a particular difference image. Generally though, darker shades of grey indicated negative values: these were the areas in which the filtered dataset was higher than the reference. Similarly, lighter shades of grey represented positive difference values which corresponded with those areas in which the filtered dataset was lower than the reference. The descriptive statistics for the resultant image differences are presented in Tables 8.9 and 8.10 (page 282).

Figure 8.61 shows the difference image resulting from the subtraction of the bilinear geometrically filtered DTM (2m resolution) (Figure 8.59) from the reference DTM (Figure 8.60). Negative differences between the two can be seen in the region of the higher vegetation (in the SE corner of the study area) and in the central area. In these regions the filtered dataset is higher than the reference model suggesting that not all of the 'above ground' points had been removed during filtering. Positive differences were minimal, other than the small area in the NW of the image. The descriptive statistics for this image indicated that the actual magnitudes of error were smaller than might be thought from observing the image. The image difference statistics revealed low standard deviation and RMSE values, in additional the mean and mode scores deviated very little from a value of zero. All of this indicates that there was very little overall difference between the height values in the geometrically filtered dataset and the reference model.

Figure 8.59: The filtered geometric linear 2m surface

Figure 8.60: The photogrammetrically derived reference model

Figure 8.61: The difference image for the geometric linear 2m dataset. Created by subtracting the Geometric filtered bilinear (2m) dataset from the photogrammetric reference DTM. The height differences in the image range from -14.87 to + 86.65 (m)(Figure 8.60)

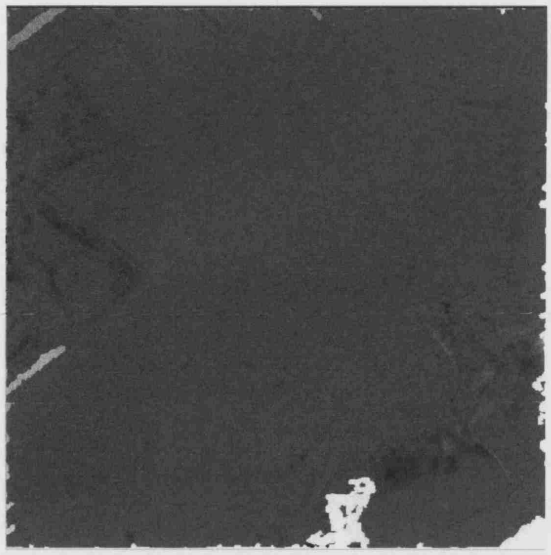

Figure 8.62: The difference image for the ATM linear 2m dataset

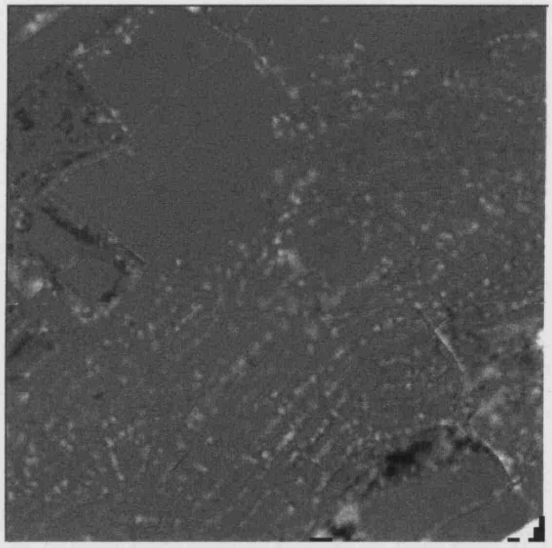

Figure 8.63: The difference image for the IRI cubic 2m dataset

Figure 8.62 shows the difference image resulting from the subtraction of the dataset filtered with the ATM algorithm. Despite the general lack of contrast in the grey tones in this image it is possible to note a general predominance of negative differences in vegetated areas. In addition to this, the ATM method generally appeared to be scouring the terrain. An example of this scouring is shown in Figure 8.62, in which the whole area of vegetation in the SE corner of the study area has been extensively degraded by the filtering process. The descriptive statistics corroborated these observations and showed the ATM method performed relatively poorly in terms of modelling bare earth heights. Indeed, this method exhibited much higher mean and mode values than the other two algorithms, and the variance of the differences was much greater, indicated by the relatively high standard deviation and RMSE values. The relative lack of success of height modelling for the ATM method can be contrasted with its success in terms of classification accuracy - this dichotomy is interesting and reflects the need for the suitability of any algorithm to be determined by the needs of the user. Some of the differences in accuracy may be explicable in terms of the reformatting procedure which was required for outputting of the ATM datasets as continuous surfaces. This procedure may have introduced an (unquantifiable) amount of error into the dataset: however it was not thought

that this process could account for all differences between the ATM datasets and the reference model.

The image difference for the IRI method (Figure 8.63) showed a different pattern from those noted for the geometric and ATM methods. In the IRI difference image, differences were observed over both buildings and vegetation. Figure 8.63 shows that there is a generally patterning to the distribution of the differences, with those in vegetated areas tending to be negative and those over building edges being positive. This suggests that in the vegetated areas the algorithm was failing to completely remove the vegetation, whilst in the built areas the bare earth surface was scoured. This scouring becomes more obvious by examination of the profile of the filtered datasets, an example of which is shown in Figure 8.64, in which both the general irregularity of the surface form and the excessive scouring in building regions. This surface may be compared with the profile of the reference DTM (Figure 8.65) which is clearly a much smoother surface. The differences between the two become even more apparent in Figures 8.66 and 8.67 which show close ups of the same part of one area. Despite these differences, the modelling of bare earth height by the IRI algorithm was relatively successful according to the descriptive statistics as indicated by the low standard deviation and RMSE scores (see Table 8.9, page 282). As with the geometric algorithm, the IRI method produced mean and modal values close to zero indicating a low overall difference between it and the reference DTM.

In terms of differences between the input datasets used for filtering, there was no substantial improvement in height accuracy offered by using the raw data. In fact, the accuracy of the raw data was lower than that produced by the nearest neighbour datasets for both the ATM and the IRI algorithms. The successes produced by the algorithms for the nearest neighbour data were clear, with the statistics for all algorithms showing an increase in accuracy. For the geometric algorithm the RMSE for the filtered nearest neighbour dataset was 0.73m, whilst that for the cubic dataset was some 2.94m. Thus, in terms of height modelling the dataset which produced the highest overall accuracy was the nearest neighbour. These conclusions mirrored

those from the classification accuracy experiment. In terms of the most accurate algorithm for height modelling, the results presented here suggested both the IRI and the geometric algorithms produced the highest accuracies.

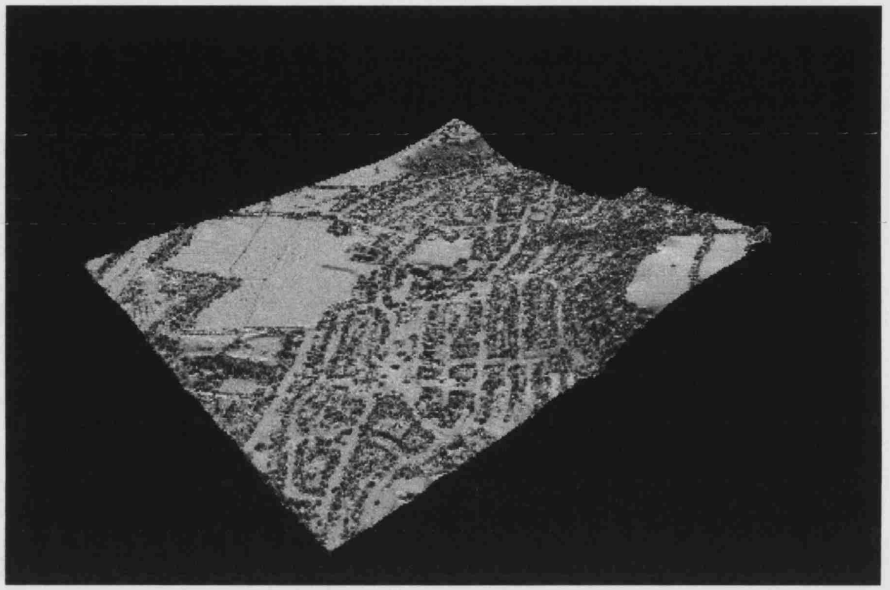

Figure 8.64: Profile of the filtered cubic 0.5m dataset using the IRI algorithm. Note the presence of some above ground objects remaining within the surface

8.5.1.1 Statistical Analysis

Some additional statistical analysis was performed in order to ascertain whether there were significant differences between the heights within the DTMs. This analysis involved comparing the mean heights in each of the datasets with those of the reference model. Whilst such analysis ignores spatial information, including the spatial variation of the mean across the study area, it does provide at least an indication of differences which can be used in further analysis. The problems associated with using spatial statistics in this study were highlighted in Chapter 4.

A student t-test was used to compare the mean heights, the formula for which is shown below (Equation 8.5). The values in each of the compared DTMs were related to each other in terms of location, and for this reason a paired t-test was calculated. One of the pre-requisites for using the t-test is that the differences

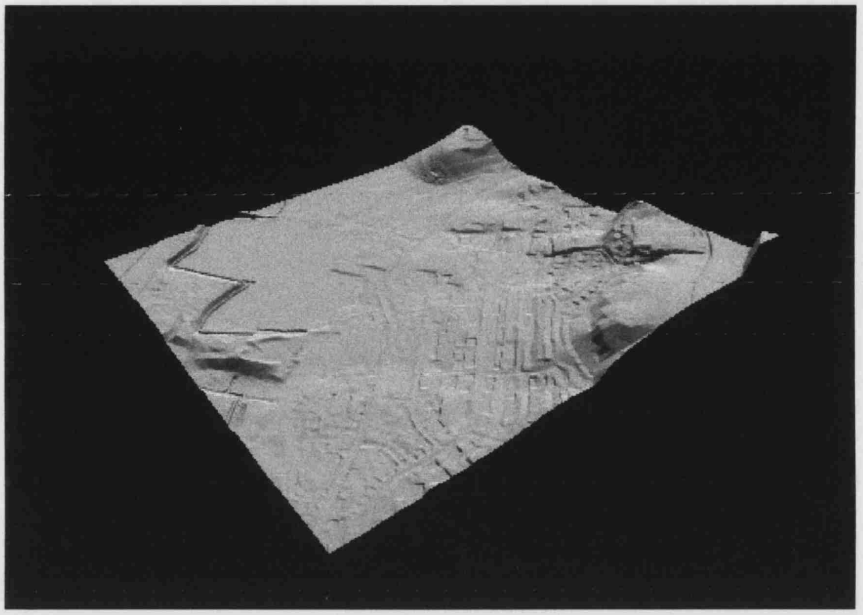

Figure 8.65: Profile of the reference DTM. Note the much smoother surface than that produced by the IRI algorithm, with almost no above ground features remaining

Figure 8.66: Zoomed in section of profile of the filtered cubic 0.5m dataset. Note the irregular surface form and the scouring in built regions

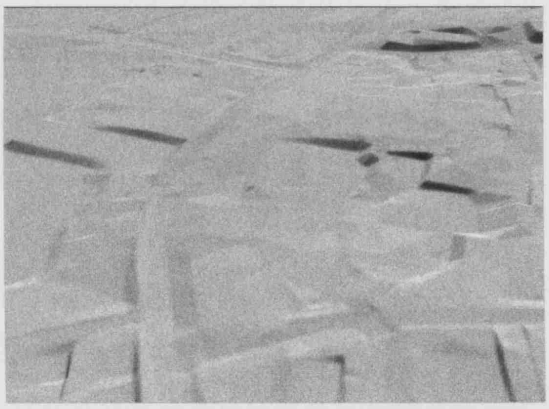

Figure 8.67: Zoomed in section of profile of the reference DTM. Note the much smoother surface in which no scouring is evident

Table 8.9: Descriptive Statistics for Image Differencing: Linear and Cubic results

		Linear			Cubic		
		0.50	**1M**	**2M**	**0.50**	**1M**	**2M**
ATM Default	MIN	-13.04	86.65	86.65	-9.09	-5.95	-7.61
	MAX	86.65	-11.35	-11.38	111.53	86.65	167.67
	MEAN	1.90	0.82	0.44	3.66	1.04	8.55
	MODE	-0.19	-0.25	-0.28	-0.14	-0.17	-0.08
	STD DEV	9.97	7.14	6.14	12.52	7.19	16.23
	RMSE	10.15	7.19	6.16	13.04	7.27	18.35
ATM Adapted	MIN	-1.75	-2.22	-4.87	-10.18	-1.23	-5.05
	MAX	86.65	86.65	86.65	86.65	86.65	86.65
	MEAN	3.08	2.81	2.58	15.68	7.69	2.55
	MODE	-0.03	-0.13	-0.22	-0.35	-0.20	-0.03
	STD DEV	11.80	11.00	10.58	15.86	17.32	10.50
	RMSE	12.20	11.35	10.89	22.30	18.95	10.81
IRI Default	MIN	-7.75	-7.75	-7.10	-6.11	-7.21	-6.14
	MAX	86.65	86.65	86.65	86.65	86.65	86.65
	MEAN	-0.17	-0.06	0.03	0.46	0.34	0.28
	MODE	-0.37	-0.38	-0.14	-0.31	-0.25	-0.34
	STD DEV	2.18	2.15	2.92	2.31	2.24	2.93
	RMSE	2.19	2.15	2.92	2.35	2.26	2.95
IRI Adapted	MIN	-9.66	-10.27	-9.68	-9.24	-7.84	-9.35
	MAX	86.65	86.65	86.65	86.65	86.65	86.65
	MEAN	-0.23	-0.10	0.00	0.46	0.33	0.25
	MODE	-0.25	-0.43	-0.27	-0.25	-0.09	-0.35
	STD DEV	2.18	2.18	2.94	2.32	2.26	2.94
	RMSE	2.19	2.18	2.94	2.36	2.29	2.95
Geometric Default	MIN	-20.40	-16.21	86.65	-23.42	-17.47	-14.69
	MAX	86.61	85.89	-14.88	85.89	85.98	86.65
	MEAN	-0.43	0.02	0.39	-0.03	0.35	0.77
	MODE	-0.34	-0.26	-0.20	0.06	-0.89	-0.04
	STD DEV	2.31	1.87	2.92	2.55	2.70	3.29
	RMSE	2.35	1.87	2.94	2.55	2.72	3.38
Geometric Adapted	MIN	-19.09	-14.67	-19.20	-19.64	-18.37	-12.67
	MAX	85.89	85.89	86.65	85.89	85.89	86.65
	MEAN	-0.53	0.04	0.52	-0.09	0.31	0.66
	MODE	-0.23	-0.14	-0.18	-0.27	-0.04	0.14
	STD DEV	2.52	2.01	3.19	2.65	2.13	3.08
	RMSE	2.57	2.01	3.23	2.65	2.15	3.15

Table 8.10: Descriptive Statistics for Image Differencing: Nearest Neighbour and raw data results

		Nearest Neighbour			Raw Data
		0.50	1M	2M	
ATM Default	MIN	-6.26	-3.45	-4.85	-5.49
	MAX	86.65	86.65	86.65	86.65
	MEAN	1.18	0.70	0.53	0.61
	MODE	-0.09	-0.29	-0.20	-0.09
	STD DEV	7.90	6.43	5.99	6.21
	RMSE	7.99	6.47	6.02	6.24
ATM Adapted	MIN	-35.89	-1.58	-4.31	-1.63
	MAX	86.65	86.65	86.65	86.65
	MEAN	10.38	2.32	1.90	2.25
	MODE	-0.47	-0.20	-0.04	-0.25
	STD DEV	15.05	10.56	9.57	10.42
	RMSE	18.28	10.82	9.76	10.66
IRI Default	MIN	-7.08	-5.63	-6.80	-5.95
	MAX	18.48	17.71	16.63	86.65
	MEAN	-0.07	-0.04	-0.01	0.00
	MODE	-0.09	-0.07	-0.03	-0.16
	STD DEV	0.70	0.69	0.66	2.13
	RMSE	0.70	0.69	0.66	2.13
IRI Adapted	MIN	-13.22	-8.16	-10.09	-9.39
	MAX	15.71	17.79	16.63	86.65
	MEAN	-0.11	-0.06	-0.03	-0.02
	MODE	-0.11	-0.05	-0.01	-0.39
	STD DEV	0.78	0.75	0.73	2.15
	RMSE	0.79	0.75	0.73	2.15
Geometric Default	MIN	-18.82	-16.89	-14.84	
	MAX	84.57	86.02	17.25	n/a
	MEAN	-0.33	0.07	0.30	
	MODE	-0.24	-0.01	0.08	
	STD DEV	1.86	2.16	0.92	
	RMSE	1.89	2.17	0.97	
Geometric Adapted	MIN	-19.14	-15.05	-12.67	
	MAX	84.57	54.33	86.65	n/a
	MEAN	-0.38	0.01	0.66	
	MODE	-0.10	-0.14	-0.14	
	STD DEV	2.16	1.08	3.08	
	RMSE	2.19	1.08	3.15	

should be approximately normally distributed, and this was confirmed by plotting histograms of the heights within the filtered datasets.

$$Calc - t = \frac{\overline{d}}{\frac{s_d}{\sqrt{n}}} \tag{8.5}$$

where: $\overline{d}$ is the mean of the paired differences, s_d is the standard deviation of the differences, and n is the number of paired differences.

The critical value of t $(Tab - t)$ was 1.96 for $\alpha = 0.025$ since this was a two sided test, and $v = (n - 1) = 4004001$ which in the tables equates to ∞. The full results of the t-test comparison are presented in Appendix K.

The results indicated significant differences between the heights within all of the laser models and the reference DTM. Clearly, although this is only a simple analysis, it does highlight the fact that whichever method of post-processing is used, some significant differences between the model and the reference data will occur.

8.5.2 Results and Discussion: Comparison with GCPs

The descriptive statistics for the residuals calculated for the differences between the GPS values and the filtered datasets are shown below in Tables 8.11 to 8.13. In the tables a value of 'unknown' is recorded if there was NODATA value in the DTM at the location of the GCP.

Table 8.11: Descriptive Statistics for GPS Comparison: Geometric Algorithm

Default	**linear**			**cubic**			**nn**		
	2m	1m	0.5m	2m	1m	0.5m	2m	1m	0.5m
min	0.04	0.00	0.00	0.04	0.02	0.01	-0.19	0.03	0.02
max	0.21	0.24	0.13	0.22	0.68	0.20	0.36	0.25	0.25
mean	0.12	0.10	0.05	0.13	0.21	0.09	0.09	0.11	0.10
std dev	0.07	0.09	0.05	0.08	0.26	0.09	0.19	0.09	0.09
RMSE	0.14	0.14	0.07	0.16	0.33	0.12	0.21	0.14	0.13
Adapted	**linear**			**cubic**			**nn**		
	2m	1m	0.5m	2m	1m	0.5m	2m	1m	0.5m
min	0.04	0.00	0.00	0.04	0.02	0.01	-0.15	0.03	0.00
max	0.33	0.24	0.13	0.46	0.68	0.20	0.36	0.25	0.19
mean	0.15	0.10	0.05	0.18	0.21	0.09	0.09	0.11	0.08
std dev	0.12	0.09	0.05	0.17	0.26	0.09	0.18	0.09	0.08
RMSE	0.19	0.14	0.07	0.25	0.33	0.12	0.20	0.14	0.11

Table 8.12: Descriptive Statistics for GPS Comparison: ATM Algorithm

Default	**linear**			**cubic**			**nn**			
	2m	1m	0.5m	2m	1m	0.5m	2m	1m	0.5m	**Raw**
min	-0.23	-0.27	-0.20	-0.10	-0.28	-0.20	-0.08	-0.20	-0.08	-0.27
max	0.07	0.07	0.04	1.62	0.07	0.06	0.13	0.08	0.11	0.07
mean	-0.03	-0.04	-0.04	0.31	-0.04	-0.04	0.03	0.00	0.03	-0.03
std dev	0.12	0.12	0.09	0.73	0.13	0.09	0.07	0.10	0.08	0.12
RMSE	0.12	0.13	0.10	0.80	0.13	0.10	0.08	0.10	0.08	0.12
Adapted	**linear**			**cubic**			**nn**			
	2m	1m	0.5m	2m	1m	0.5m	2m	1m	0.5m	**Raw**
min	-0.26	-0.27	-0.18	-0.24	-0.28	unknown	-0.08	-0.01	unknown	-0.22
max	0.07	0.07	0.05	0.06	0.07	unknown	0.13	0.08	unknown	0.07
mean	-0.04	-0.04	-0.03	-0.04	-0.04	unknown	0.03	0.02	unknown	-0.02
std dev	0.12	0.12	0.08	0.11	0.13	unknown	0.07	0.03	unknown	0.10
RMSE	0.13	0.13	0.09	0.12	0.13	unknown	0.08	0.04	unknown	0.10

Table 8.13: Descriptive Statistics for GPS Comparison: IRI Algorithm

Default	**linear**			**cubic**			**nn**			
	2m	1m	0.5m	2m	1m	0.5m	2m	1m	0.5m	**Raw**
min	-0.20	-0.09	-0.10	-0.23	-0.14	-0.10	-0.07	-0.05	-0.08	unknown
max	0.03	0.03	0.01	0.04	0.02	0.01	0.01	0.04	0.01	unknown
mean	-0.05	-0.02	-0.02	-0.05	-0.03	-0.02	-0.01	-0.01	-0.02	unknown
std dev	0.08	0.04	0.04	0.10	0.06	0.04	0.03	0.03	0.03	unknown
RMSE	0.09	0.04	0.05	0.11	0.06	0.05	0.03	0.03	0.04	unknown

Adapted	**linear**			**cubic**			**nn**			
	2m	1m	0.5m	2m	1m	0.5m	2m	1m	0.5m	**Raw**
min	-0.22	-0.20	-0.19	-0.23	-0.42	-0.10	-0.07	-0.05	-0.08	unknown
max	0.04	0.06	0.03	0.05	0.01	0.02	0.01	0.04	0.01	unknown
mean	-0.07	-0.04	-0.04	-0.05	-0.10	-0.02	-0.01	-0.01	-0.02	unknown
std dev	0.09	0.09	0.09	0.10	0.16	0.04	0.03	0.03	0.03	unknown
RMSE	0.11	0.10	0.10	0.11	0.19	0.05	0.03	0.03	0.04	unknown

The results show there are many similarities in the results of the comparisons for the three algorithms, namely that the nearest neighbour dataset produced smaller height errors than other datasets. It should be noted that for the geometric algorithm there was, in fact, very little difference between the errors produced by either the linear or the nearest neighbour datasets. From the tables it is also clear that the cubic dataset again contained much greater differences between the modelled surface and the RTK GPS points than any of the other datasets. The results for the IRI comparison produced some extremely small differences between the nearest neighbour dataset and the GPS points. This was slightly surprising, for two reasons. Firstly, the image differencing and classification accuracy investigations had shown that this algorithm had not been particularly successful and it had therefore been expected that the algorithm may also have demonstrated problems for the modelling of the bare earth. Secondly the results of the capture error quantification (Chapter 5) indicated that the mean data capture error for the laser points was 16.34cm (0.16m) - the difference results for the IRI algorithm here suggest that the errors in the dataset are in fact much lower than this in bare earth areas and in fact lay within the error range of the GPS equipment (which was 4.5cm). Thus, whilst the IRI algorithm was found to perform poorly in the removal of above ground features, and the modelling of the surface across these, in bare earth areas it was found to introduce the least error into the height models. These results are important, as they demonstrate that the suitability of any algorithm is wholly dependent on the needs of the user.

No statistical analysis was conducted on the results of the GPS comparison simply because the number of GPS points in the sample available was not sufficient to provide a representative sample of the whole area.

8.6 Answering the Research Questions

In terms of attaining the aims of the thesis as established in Chapter 1, the results presented above need to be directly compared with one another in order to establish whether there are statistically significant differences between the results. The results of these tests are presented below with specific reference to the research question which they address. In this section statistical differences between the methods are tested for with reference to the classification accuracies.

8.6.1 Q1 Are there significant differences between the classification accuracies for different filtering methods?

The results from this test are perhaps the most important for users defining their post-processing methodology, as the choice of filtering method is the most common decision which users have to make in their post-processing. All users wishing to create a DTM from discrete laser points will have to make a decision regarding which method they are to use, finer details such as whether to alter the filtering parameters are only relevant once the filtering algorithm has been chosen, and not all users will wish to do this.

In order to ascertain if there were statistically significant differences in the classification accuracies in the filtered datasets from different algorithms, the contingency tables for each dataset were again compared with all others using the z-test.

Table 8.14 is a matrix summarising the results for each of contingency table comparisons. For each comparison there were between nine and ten individual z-test scores - one for each of the datasets compared. For example the comparison of the IRI default method with the ATM default method produced ten results as each spatial interpolation method at each grid size was compared for the IRI method was compared with its ATM method counterpart. All of the results are presented in

Appendix L: however to aid discussion here these results were simplified by assigning a summary symbol, of either - or *, to represent the overall results from each comparison. A dash (-) symbol indicated that overall the majority of compared datasets did not display a significant difference, while a star symbol (*) indicated that overall the datasets showed that there was a significant difference between the methods compared. This simplification is used purely to present the general characteristics in the results. Further discussion of some of the finer points is made in the text below.

Table 8.14: Summary of Results

Filtering Method	IRI default	IRI adapted	ATM default	ATM adapted
ATM default	-	-		
ATM adapted	-	-		
Geometric default	-	*	-	*
Geometric adapted	*	*	-	*

Table 8.14 shows that in some instances there was a significant difference in classification accuracy produced by using different filtering algorithms. This finding indicates that the choice of filtering algorithm can play a significant part in the introduction of error into the modelling. In particular, the geometric method was found to produce significantly different classification accuracies to either the IRI or the ATM methods. Referral to Tables 8.6 and 8.7 shows that the geometric algorithm was more accurate.

In the comparison of the adapted algorithms for the IRI and the ATM methods there was a less obvious difference - given that only five out of the ten comparisons produced a statistically significant difference. This was thought to be because the adaptations of the original methods meant that they were more suited to filtering the datasets used in this investigation. The improvements offered by these adaptations are likely to have reduced any differences produced by the original methods.

8.6.2 Q2 Is there a significant difference in accuracy caused by filter method parameter changes?

The results of this test provided an indication of the sensitivity of each the filtering algorithms. This is important from a user's perspective when choosing the most suitable filtering algorithm for a given task. Whilst in practice it may not be possible for a user to know exactly how much error they are introducing, it is important that they are aware that using different parameter settings for a particular filtering method may introduce different amounts of error, and therefore uncertainty, into their models.

In order to test for parameter sensitivity of the three algorithms, filtering was conducted using different parameter settings and significant differences between the results were tested for, again using the z-test. For example, the accuracy of the filtered dataset created using the default method in IRI was directly compared with the filtered dataset created using the modified IRI algorithm written as part of this research. A significant difference between these two would indicate that the filtering accuracy is sensitive to changes in parameter settings for this algorithm.

The results are shown in Tables 8.15 and 8.16. Table 8.15 details the calculated values of z for each comparison, and Table 8.16 summarises the results of the hyotheses.

Table 8.15: Z-test Results for Comparison of Different Settings for the Three Algorithms

Algorithm	**Dataset**				
		0.5m	1m	2m	Original
IRI	Linear	2.50	4.13	1.27	
	Cubic	5.77	3.94	2.74	
	NN	2.15	2.16	0.64	
	Raw Data				0.64
ATM	Linear	0.29	0.56	1.03	
	Cubic	5.48	3.51	0.78	
	NN	0.19	0.26	0.60	
	Raw Data				0.60
Geometric	Linear	1.852	3.7	3.6	
	Cubic	3.009	3.1	2.3	
	NN	4.001	1.7	4.5	

In Table 8.16 below, a star (*) symbol indicates the acceptance of the alternative hypothesis that there is a significant difference between the datasets at the 5% significance level for a two-tailed test, whilst a dash (-) symbol indicates the acceptance of the null hypothesis of no significant difference. This same notation is used in the remainder of the summary tables in this thesis.

Table 8.16: Hypothesis Results for Comparison of Different Settings for the Three Algorithms

Algorithm	**Dataset**				
		0.5m	1m	2m	Original
IRI	Linear	*	*	-	
	Cubic	*	*	*	
	NN	*	*	-	
	Raw Data				-
ATM	Linear	-	-	-	
	Cubic	*	*	-	
	NN	-	-	-	
	Raw Data				-
Geometric	Linear	-	*	*	
	Cubic	*	*	*	
	NN	*	-	*	

The results from this experiment showed that for both the geometric and the IRI algorithms there were significant differences in the accuracies of the classification results. Referring to Tables 8.6 and 8.8 it can be seen that the differences in the geometric method were caused by the increased accuracy offered by the adapted settings, whilst the reverse was true for the IRI method.

There was no significant difference between the settings for the ATM method, which produced consistently high classification accuracies. In this respect the ATM was a more robust algorithm as it was not as sensitive to changing parameters as the other methods tested. As such it may be more attractive to those users wishing to use the same parameter settings for post-processing.

8.6.3 Q3 Is there a significant difference in the classification accuracy between the gridded datasets and the raw data?

The contingency tables for the raw and interpolated datasets were compared for both the IRI and the ATM methods. The results from this test only pertain to the ATM and the IRI methods as these were the only methods which allowed both raw and interpolated data to be filtered. The results are shown below in Table 8.17, from which it can be seen that for the majority of datasets there is no significant difference in the classification accuracies of the interpolated and the raw data. Despite this overall tendency, there are some significant differences between the datasets, most notably the comparison between the classification accuracies of the raw data and the nearest neighbour 1m dataset for the adapted ATM method. Reference to Table 8.7 reveals that this was caused by the much higher user accuracies of the interpolated dataset. This indicates that, in this instance, by using interpolated data a higher filtering accuracy can be attained.

Results such as these are extremely important for laser data users, especially given that they refute one of the most widely held beliefs in the laser scanning community - that using raw data for modelling will **always** produce higher accuracy results. As Tables 8.6 to 8.8 indicate, clearly this is not the case.

8.6.4 Q4 Are there significant differences between the classification accuracies of the different gridded datasets?

The aim of this part of the investigation was to establish whether the method used for data restructuring affected the classification accuracy. This has already been alluded to in some of the analyses above, but required a robust statistical test to show whether there are, in fact, any significant differences.

The results from the contingency tables for each of the interpolation methods (bilinear, bicubic, and nearest neighbour) for each filtering method were compared with

Table 8.17: $Z - test$ and hypotheses results to show the effect of data structure on classification accuracy of filtering algorithms

Algorithm	Data Structuring Method	Grid space	Z-test result	Hypothesis result
IRI Default	**Bilinear**	0.5	1.92	-
		1	3.99	*
		2	2.92	*
	Bicubic	0.5	0.18	-
		1	1.73	-
		2	2.61	*
	NN	0.5	0.43	-
		1	0.55	-
		2	0.52	-
IRI Adapted	**Bilinear**	0.5	2.10	*
		1	2.63	*
		2	4.21	*
	Bicubic	0.5	3.13	*
		1	2.66	*
		2	2.25	*
	NN	0.5	0.30	-
		1	1.17	-
		2	2.63	*
ATM Default	**Bilinear**	0.5	1.53	-
		1	1.33	-
		2	1.14	-
	Bicubic	0.5	0.38	-
		1	1.66	-
		2	0.22	-
	NN	0.5	0.81	-
		1	2.14	*
		2	1.43	-
ATM Adapted	**Bilinear**	0.5	0.79	-
		1	0.87	-
		2	0.93	-
	Bicubic	0.5	6.16	*
		1	2.97	*
		2	0.49	-
	NN	0.5	0.06	-
		1	0.87	-
		2	1.01	-

each other to test for significant differences in classification accuracy. As with the investigations above, a z-test was used to test for significant difference. The z-test scores and hypotheses results are shown in Tables 8.18 to 8.23.

Table 8.18: Z-test and hypotheses results for IRI default algorithm

IRI default method	**Datasets compared**	**0.5m**	**1m**	**2m**
Z-test	Bilinear/bicubic	2.12	5.49	5.47
	Bilinear/nearest neighbour	2.34	3.52	2.42
	Bicubic/nearest neighbour	0.25	2.29	3.14
	Datasets compared	**0.5m**	**1m**	**2m**
Hypothesis Result	Bilinear/bicubic	*	*	*
	Bilinear/nearest neighbour	*	*	*
	Bicubic/nearest neighbour	-	*	*

Table 8.19: Z-test and hypotheses results for adapted IRI algorithm

IRI adapted method	**Datasets compared**	**0.5m**	**1m**	**2m**
Z-test	Bilinear/bicubic	5.08	5.25	6.29
	Bilinear/nearest neighbour	1.84	1.45	1.68
	Bicubic/nearest neighbour	3.48	3.79	4.82
	Datasets compared	**0.5m**	**1m**	**2m**
Hypothesis Result	Bilinear/bicubic	*	*	*
	Bilinear/nearest neighbour	-	-	-
	Bicubic/nearest neighbour	*	*	*

Some significant differences between the classification accuracies were noted, indicating that the choice of interpolation method can significantly affect the quality of the filtered DTM. The main findings from the z-test investigation are summarised below.

- For the IRI algorithm there were significant differences between most of the datasets, indicative of a lack of consistency in classification accuracies. There was no significant difference between the bilinear and the nearest neighbour datasets for the adapted settings. In general, this observation was also repeated for both the ATM and the geometric methods. Indeed, Tables 8.6 to

Table 8.20: Z-test and hypotheses results for default ATM algorithm

ATM default method	Datasets compared	0.5m	1m	2m
	Bilinear/bicubic	1.15	0.35	1.35
Z-test	Bilinear/nearest neighbour	0.75	0.81	0.28
	Bicubic/nearest neighbour	0.42	0.46	1.64
	Datasets compared	**0.5m**	**1m**	**2m**
	Bilinear/bicubic	-	-	-
Hypothesis Result	Bilinear/nearest neighbour	-	-	-
	Bicubic/nearest neighbour	-	-	-

Table 8.21: Z-test and hypotheses results for adapted ATM algorithm

ATM adapted method	Datasets compared	0.5m	1m	2m
	Bilinear/bicubic	6.82	3.82	0.45
Z-test	Bilinear/nearest neighbour	0.85	0.00	1.93
	Bicubic/nearest neighbour	6.13	3.82	1.50
	Datasets compared	**0.5m**	**1m**	**2m**
	Bilinear/bicubic	*	*	-
Hypothesis Result	Bilinear/nearest neighbour	-	-	-
	Bicubic/nearest neighbour	*	*	-

Table 8.22: Z-test and hypotheses results for default Geometric algorithm

Geometric default method	Datasets compared	0.5m	1m	2m
	Bilinear/bicubic	0.48	1.08	0.82
Z-test	Bilinear/nearest neighbour	1.34	3.52	0.90
	Bicubic/nearest neighbour	1.83	2.49	0.09
	Datasets compared	**0.5m**	**1m**	**2m**
	Bilinear/bicubic	-	-	-
Hypothesis Result	Bilinear/nearest neighbour	-	*	-
	Bicubic/nearest neighbour	-	*	-

Table 8.23: Z-test and hypotheses results for adapted Geometric algorithm

Geometric adapted method	Datasets compared	0.5m	1m	2m
Z-test	Bilinear/bicubic	0.48	1.08	0.82
	Bilinear/nearest neighbour	1.34	3.52	0.90
	Bicubic/nearest neighbour	1.83	2.49	0.09
	Datasets compared	**0.5m**	**1m**	**2m**
Hypothesis Result	Bilinear/bicubic	-	-	-
	Bilinear/nearest neighbour	-	*	-
	Bicubic/nearest neighbour	-	*	-

8.8 show that both the bilinear and the nearest neighbour algorithms consistently produced the highest accuracies.

On the basis of these findings it is suggested that users wishing to reduce error in their filtered models should use these interpolation methods when restructuring their raw laser data.

- For the ATM algorithm there were no significant differences between the datasets for the default parameter settings, suggesting that this algorithm is much less sensitive to input surface form or parameter changes. For this reason it was considered to be the most robust of all the algorithms tested within this investigation.

- Finally, the geometric algorithm showed relatively little significant difference between the majority of datasets, indicating that this algorithm was little influenced by surface changes.

8.6.5 Q5 Does the spatial resolution of the data affect the characteristics of the error in the filtered datasets?

To test the effect of spatial resolution on classification accuracy the contingency tables were analysed again for statistically significant differences. Full results from this investigation are presented in Appendix M. For simplicity only one of the results

tables is presented here. Table 8.24 shows the z-test and hypotheses results for the comparison of the different spatial resolutions for the ATM algorithm. From the results in Table 8.24 and Appendix M it would appear that in terms of classification accuracy there is generally no significant difference caused by changes in the spatial resolution of input datasets. Thus, whilst it was shown in Chapter 7 that changes in spatial resolution may affect the quality of DSMs, the results here would suggest that in fact spatial resolution is not an influencing factor in terms of DTM accuracy - at the resolutions examined here (i.e. those similar to the point spacing of the input laser data).

The implication of this finding is that data users who wish to create a DTM can choose a low spatial resolution input dataset and so reduce processing time and data volume without compromising the accuracy of the resultant DTM. This finding has many practical implications for laser data users, particularly given the fact that laser datasets are often very large and methods for reducing data volumes without reducing data accuracy are highly sought after.

Table 8.24: Comparison of ATM default results

Dataset	Comparison	Z-test result	Hypothesis result
Bilinear	0.5m vs 1m	0.20	-
	0.5m vs 2m	0.38	-
	1m vs 2m	0.18	-
Bicubic	0.5m vs 1m	1.29	-
	0.5m vs 2m	0.59	-
	1m vs 2m	1.87	-
Nearest	0.5m vs 1m	1.37	-
Neighbour	0.5m vs 2m	0.64	-
	1m vs 2m	0.72	-

8.7 Conclusions and the Significance of the Findings

Many of the conclusions from the investigations presented in this chapter will be of practical use to ALS data users, and could easily be used within a simple decision support system in order to reduce errors introduced during data post-processing. The main conclusions are summarised below.

1. The choice of filtering algorithm can significantly affect the amount of error within a resultant dataset. In terms of overall accuracy, the accuracy of a filtering classification could be raised by 30% by using the geometric algorithm in place of the IRI method.

2. The geometric algorithm was found to produce the highest overall classification accuracies, and the highest accuracy for height modelling.

3. For all of the algorithms, the lowest errors (both in terms of classification and modelled heights) were produced by using restructured data rather than the raw data. The datasets restructured using the nearest neighbour algorithm produced the least error. In other words, the use of raw data (i.e. that held in its original irregular data structure) does not necessarily produce higher accuracy filtered datasets, even when using a filtering algorithm tailored for use with irregular data. It is a widely held belief in the laser scanning community that one should always use data in its raw format: however, the results of this investigation show that using either bilinear or nearest neighbour interpolation algorithms to restructure the original data can improve the classification accuracy in the filtering stage by up to 21% (based on comparison of bilinear 2m overall classification with raw data overall accuracy for the IRI adapted algorithm - see Table 8.6). These results are based on the analysis of the three filtering algorithms used within the context of this investigation: further work should investigate the transferability of these observations to other filtering algorithms in order to confirm (or otherwise) the observations.

4. Spatial resolution of the input points was found not to significantly affect error,

meaning that in order to reduce processing times and data volumes users may use lower resolution data (for example 2m rather than 0.5m resolution) without an associated increase in filtering error. However, it should be noted here that there is likely to be a limit to this finding. Based on this investigation the finding relates to grid resolutions of order $\pm$ 2 times smaller or larger than the nominal raw laser data spacing.

5. There are clear differences in the spatial patterns of both Type I and Type II errors created by the different algorithms. The suitability of the algorithm therefore largely depends upon which type, and in which areas, the user wishes to minimise error.

8.8 Chapter Summary

In this chapter, three filtering algorithms have been compared to analyse the difference in both the magnitudes and the spatial patterns of error which were introduced during post-processing. The effect of using different input data, different spatial resolutions and different parameter settings was investigated. The results were surprising, and were perhaps counter-intuitive given the results from the investigation presented in Chapters 6 and 7. The following chapter investigates the errors introduced during the subsequent segmentation stage of processing, which is closely related to the filtering process.

Chapter 9

Analysis of the Error Arising Through Segmentation

9.1 Chapter Introduction

Segmentation is the process of grouping neighbouring laser points, or pixels, with common characteristics. These groups are usually subsequently classed as one of a range of 'above ground' classes. This technique is used in the identification of buildings, trees and other surficial terrain objects, and directly follows the filtering stage in the processing of data. The segmentation process is mandatory for the reconstruction of objects for inclusion in 3D city models. For this reason most studies have focused on classifying objects in urban areas, with particular focus on candidate building points (as shown in Chapter 3). However, very little research has been conducted on assessing the errors introduced during segmentation, and no published reports exist comparing the segmentation results of different algorithms.

Information about the accuracies of both the classifications and the resultant modelled heights of features is especially important for those users who wish to create 3D objects from the classified points. Methods for reconstructing objects are common

in the literature (e.g. Matikainen et al. 2003, Vosselman 2003) and many of these present complex algorithms which aim to produce high accuracy results. This accuracy is largely dependent on the success of the segmentation procedure. Information about this success could also help users to decide which segmentation algorithm is most appropriate to their needs, particularly if this accuracy is related to particular classes or if the errors are spatially dependent.

9.2 Chapter Aims

This chapter aims to provide information regarding the characteristics of errors which may be introduced during segmentation of ALS datasets. Two segmentation methods are investigated and directly compared. Results show whether and how factors such as segmentation method, parameter settings and input data structure can affect the classification and height accuracies. In effect, this investigation is similar to that of the filtering study presented in the previous chapter, with the emphasis on this chapter being on the 'above ground' points rather than those classified as bare earth.

In terms of quantifying the post-processing error budget (Equation 2.13) this chapter presents the quantified segmentation errors (ε_s) in the post-processing workflow.

9.3 Segmentation methods

Two segmentation methods were used in this investigation: the geometric algorithm and the ATM algorithm. These methods can be used both for filtering (as seen in Chapter 8) and segmentation. Emphasis in this section is upon explaining how the two algorithms differentiate between two classes of 'above ground' objects which are common in the urban environment: buildings and vegetation. As with the filtering investigation, both segmentation algorithms are used with two different settings

(default and adapted) in order to test for the stability of the methods.

9.3.1 Geometric Segmentation

Segmentation in the geometric algorithm is based on differentiating between the characteristics of the agglomerated polygons. As with the filtering procedure, for the segmentation the image polygons are assigned a value depending upon their relationship with neighbouring polygons or according to their geometric shape. The settings for the classification rule base, and the weighting for the class membership functions are described below. The settings were the same for both the default and the adapted methods, although different weights were used for defining the initial image object characteristics (as explained in Chapter 8).

The process for the classification of image objects employed a fuzzy rule base. This approach involves the combination of one or more conditions by the operator. In order to include features into a fuzzy rule base, membership functions for each of the features (or classes) have to be defined. The membership functions are, effectively, summaries of known information about a certain class. For example, one class may be characterised by low mean values in a certain image layer. This information can be used to classify the image. The graphical membership functions shown in Figures 9.2 to 9.5 are examples of those used in the classification. The membership functions are defined by their left and right border values, plus the function slope.

The bare earth class was defined based on the calculated values for each of the image objects for the mean difference to neighbouring image objects. An image indicating the calculated mean difference values for each of the objects is shown in Figure 9.1. In terms of the membership thresholds for this classification, values ranged between -10 and 2m. This is shown in Figure 9.2. The building class was defined in a similar way and also used the calculated values of the mean difference to neighbours: the membership function for this class is shown in Figure 9.3. The vegetation class was defined in a slightly different way, and used the calculated values of the elliptical fit

of the image objects (the calculated values are shown in Figure 9.4). In effect, these values were derived by calculating the shape of the image objects and assessing to what degree they presented elliptical shapes. The vegetation image objects had a much higher elliptical fit and were distinguished from the other classes in this way. The membership function for the vegetation class is shown in Figure 9.5.

Figure 9.1: Showing the relative values of the image objects (for the whole Bristol 1 km^2 study area) for the mean difference to neighbours calculation. Values range between -7 and 16.36m: the darker the pixel the lower the mean difference to neighbours value

9.3.2 ATM Segmentation

The ATM segmentation algorithm is based upon the classification of the ALS dataset solely on the heights of the points above ground. The ground surface had been defined from the bare earth classified points created during filtering. In this respect it was assumed that the height values alone could be used to characterise certain land covers. For both the default and the adapted settings there were three classes required to produce the final binary (buildings and vegetation) classification: the

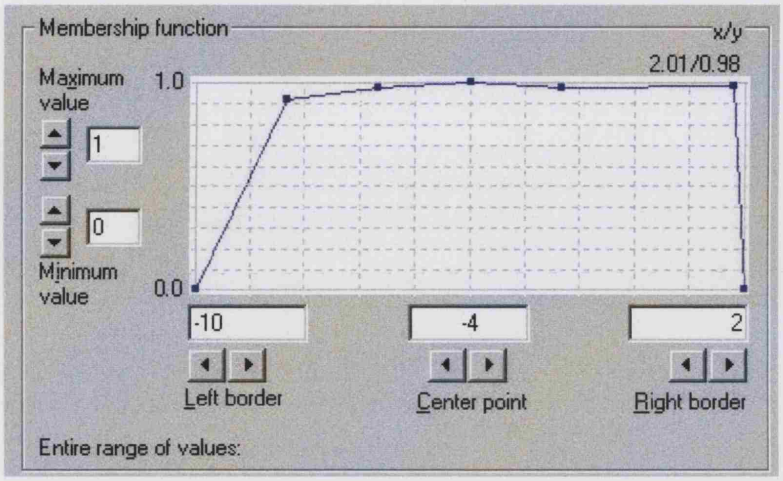

Figure 9.2: Showing the membership function for the bare earth classification. Values defining the threshold ranged between 2 and -10. The left-hand axis indicated the level of fuzziness. For example, the part of the range which is fuzzy lies approximately between -10 and -5.

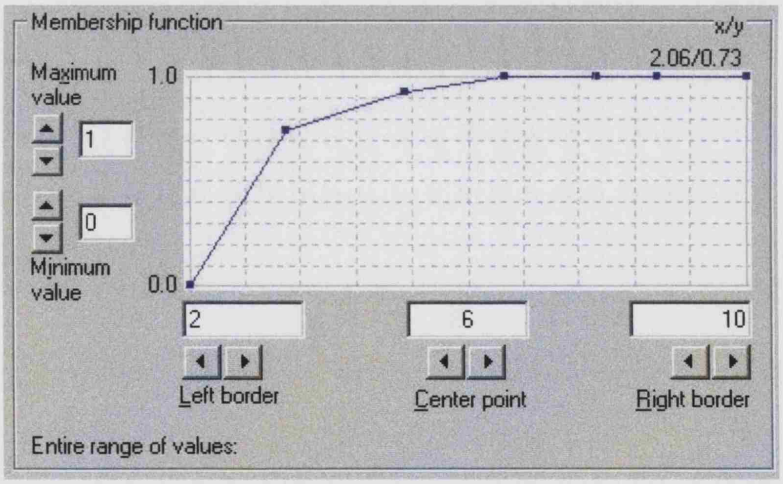

Figure 9.3: Showing the membership function for the building classification. The form of the slope indicates that lower values (between 2 and 5) are more fuzzy and are less likely to be classified definitely as building.

Figure 9.4: Showing the relative values of the image objects for the elliptical fit calculation. The calculated values range between 0.09 and 0.9, the darker the pixel the lower the elliptical fit value

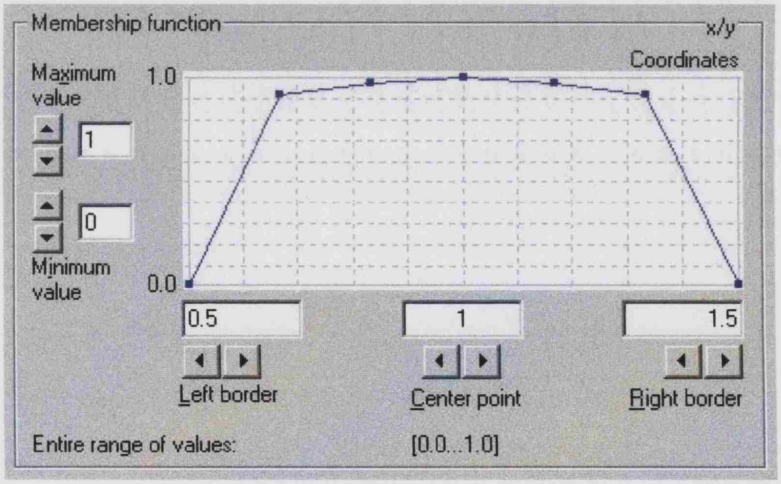

Figure 9.5: Showing the membership function for the vegetation classification. The form of the slope indicates that values of approximately 1 should be classified as vegetation, but those values greater or less than 1 are much less likely. The limits of the range of acceptable vegetation values are 0.5-1.5, beyond this range image objects will not be considered to be vegetation.

preliminary three classes were buildings, high vegetation and low vegetation. The height thresholds (for both settings of the algorithm) were as follows: low vegetation 0.5 - 3m, buildings 3.1 - 20m, high vegetation 20.1 - 150m.

The ATM algorithm does allow other values to characterise the classes, and these include classifying based on minimum elevations, by absolute elevations, and by colour, echo (multiple returns) and intensity. The dataset used in this study was most suited to a classification based on the height above ground.

The IRI algorithm, which was the other algorithm investigated in the previous chapter, could not be analysed in this way as it did not differentiate between different classes of 'above ground' objects.

9.4 Assessing the Success of the Segmentation

The errors introduced during various segmentation routines were assessed in two ways: firstly in terms of the errors within the classifications, and secondly in terms of the errors within the modelling of the classified surficial features. As with the previous investigations, both the magnitude and the spatial patterning of errors was investigated.

9.4.1 Classification Accuracy

The assessment of classification performance was again expressed as a contingency table for each dataset. An example is shown in Table 9.1 which shows the classification accuracy results for the segmentation of the linear 0.5m dataset using the geometric algorithm. In this example, the ground classification has been included for context. Whilst this study is essentially concerned with the ability of the algorithms to differentiate between buildings and vegetation, it is interesting to compare the accuracy of the ground classification for context, in order to see the relative difference

in accuracies.

Table 9.1: Contingency Table Showing Segmentation Accuracy for Linear 0.5m Dataset Using the Geometric Algorithm

		Truth data			
		Ground	Building	Vegetation	Row Total
Classification data	Ground	52	5	16	73
	Building	17	16	32	65
	Vegetation	12	38	25	75
	Column Total	81	59	73	213

Each contingency table was also used to compute producer's accuracy results to give an indication of the success of the classification for each of the individual classes. The results are shown in Tables 9.2 to 9.5 below. Further results and graphics relating to the producer's accuracy analysis for all three classes (ground, building, and vegetation) can be found in Appendix J. Two examples of the results are presented in Figures 9.6 and 9.7 as individual bar charts in order to highlight the principal features.

From the results in the tables and Figures 9.6 and 9.7, it would appear that for the ATM algorithm was more successful in classifying building points than vegetation, while the geometric algorithm showed no variation in performance between categories. There are systematic differences between the algorithms with respect to their success in predicting the three classes. The classification accuracy for the vegetation class was low, especially for the geometric algorithm which predicted as low as 30% of occurrences. The ATM algorithm was slightly more successful, predicting around 40% of vegetation occurrences, but this success rate is still lower than that for the building or ground classifications for this algorithm. For the ATM method the success rate in classifying buildings was much more successful, ranging between 75 - 80% for the different resolutions.

From the results presented in Tables 9.2 to 9.5, there does not appear to be any

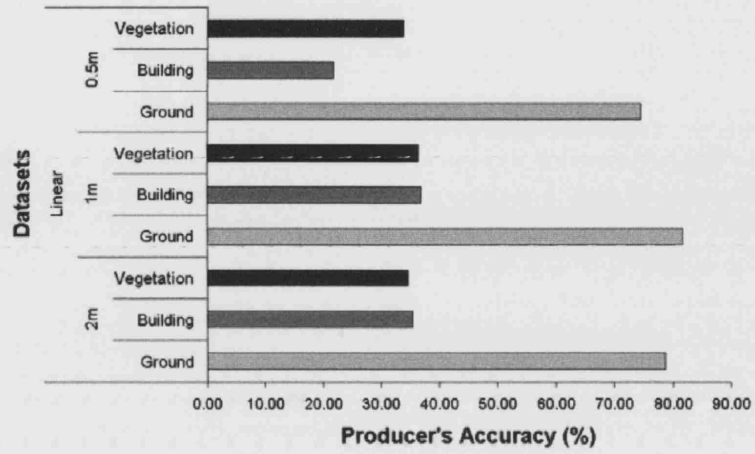

Figure 9.6: Showing producer's accuracy results for the segmentation of the linear dataset with the adapted geometric algorithm

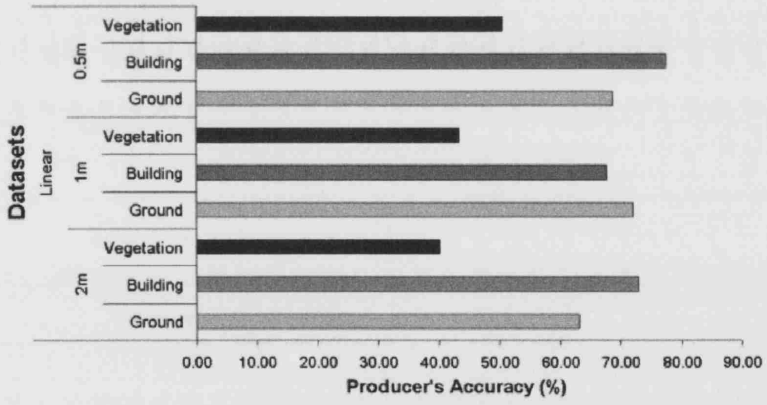

Figure 9.7: Showing producer's accuracy results for the segmentation of the linear dataset with the adapted ATM algorithm

discernable difference between the results for the different input datasets for the geometric algorithm, suggesting that the method of data restructuring did not appear to affect the classification accuracies. For the ATM algorithm a slight improvement in the accuracy of the building classification arose when using bilinearly restructured data. There was little difference between the accuracies produced by the bilinear dataset and those produced by the raw data.

Overall, the ATM algorithm appeared to produce higher accuracies than the geometric for the same datasets. In order to ascertain whether there were significant differences in the producer's accuracy results, they were analysed using the z-test. The results from this statistical investigation are presented in Section 9.5, where they are related directly to the research questions and aims of the thesis.

Table 9.2: Producer's Accuracy Results for Segmentation Using the Geometric Algorithm with Default Settings

			Producer's Accuracy %
Linear	2m	Building	51.16
		Vegetation	58.21
	1m	Building	39.74
		Vegetation	41.38
	0.5m	Building	25.71
		Vegetation	34.38
Nearest Neighbour	2m	Building	52.08
		Vegetation	30.36
	1m	Building	43.16
		Vegetation	33.33
	0.5m	Building	43.04
		Vegetation	38.83
Cubic	2m	Building	53.85
		Vegetation	43.21
	1m	Building	38.00
		Vegetation	38.24
	0.5m	Building	39.24
		Vegetation	30.59

Table 9.3: Producer's Accuracy Results for Segmentation Using the Geometric Algorithm with Adapted Settings

			Producer's Accuracy %
Linear	2m	Building	35.34
		Vegetation	34.57
	1m	Building	36.70
		Vegetation	36.23
	0.5m	Building	21.59
		Vegetation	33.70
Nearest Neighbour	2m	Building	41.18
		Vegetation	34.18
	1m	Building	42.34
		Vegetation	33.94
	0.5m	Building	36.94
		Vegetation	33.21
Cubic	2m	Building	45.26
		Vegetation	33.83
	1m	Building	40.19
		Vegetation	34.81
	0.5m	Building	29.33
		Vegetation	31.75

Table 9.4: Producer's Accuracy Results for Segmentation Using the ATM Algorithm with Default Settings

			Producer's Accuracy %
Linear	2m	Building	63.01
		Vegetation	37.84
	1m	Building	62.03
		Vegetation	39.66
	0.5m	Building	75.68
		Vegetation	55.56
Nearest Neighbour	2m	Building	55.71
		Vegetation	47.44
	1m	Building	60.82
		Vegetation	53.85
	0.5m	Building	58.23
		Vegetation	52.11
Cubic	2m	Building	54.84
		Vegetation	58.93
	1m	Building	74.29
		Vegetation	47.06
	0.5m	Building	58.33
		Vegetation	47.54
Raw data		Building	62.16
		Vegetation	43.84

Table 9.5: Producer's Accuracy Results for Segmentation Using the ATM Algorithm with Adapted Settings

			Producer's Accuracy %
Linear	2m	Building	73.02
		Vegetation	40.24
	1m	Building	67.80
		Vegetation	43.53
	0.5m	Building	77.61
		Vegetation	50.60
Nearest Neighbour	2m	Building	68.12
		Vegetation	46.67
	1m	Building	70.31
		Vegetation	53.68
	0.5m	Building	70.00
		Vegetation	47.00
Cubic	2m	Building	67.61
		Vegetation	47.37
	1m	Building	79.63
		Vegetation	35.09
	0.5m	Building	63.33
		Vegetation	50.00
Raw data		Building	74.24
		Vegetation	48.81

9.4.2 Types of Error

As with the filtering investigation, there are different types of segmentation errors, the locations of which may help users to decide which of the algorithms is the most suitable for their needs. It may also help drive future research as it identifies potential problems with the algorithm.

Two examples of the spatial patterning of segmentation errors are shown in Figures 9.8 and 9.9. Figure 9.8 shows the classified errors in the 1m cubic dataset, segmented using the ATM algorithm. There is some patterning in the occurrence of the different types of error in this example. Errors in the vegetation class are the most widespread, and appear to predominantly occur over buildings: this seems to suggest that some parameter within the algorithm currently means that it is difficult to distinguish between building and vegetation objects. This is supported by the observation that the errors in the building class tend to occur over vegetation.

The spatial characteristics of the errors in the geometric dataset (Figure 9.9) are also clearly visible. Errors in the vegetation classification can be found mainly on the edges of buildings, such as the example shown in Figure 9.10, whilst most of the bare earth errors occur on very tall vegetation, such as that shown in Figure 9.11. The errors in the building classification also display distinctive patterning, with most concentrated on areas of low lying vegetation (Figure 9.12). This patterning would suggest that the geometric segmentation algorithm has difficulty in differentiating between buildings and vegetation, more worrying still is the evident confusion between the ground and vegetation classes. This latter fault was thought to arise because the geometric algorithm searches for edges, yet in large areas of vegetation such features may only occur at the break between edge of the vegetation mass and the bare earth. In the central areas of any vegetation mass, classification errors may well occur, as there are few changes in height. Figure 9.11, which shows a close up of one of the vegetation areas containing misclassified bare earth errors, appears to confirm that this is indeed the case.

Figure 9.8: Aerial photograph of study area overlain with classification errors for the ATM cubic 1m dataset. Red crosshairs are errors in ground class, blue crosshairs show errors in building class, green crosshairs show errors in vegetation class.

Figure 9.9: Aerial photograph of study area overlain with classification errors for the geometric linear 2m dataset. Red crosshairs are errors in ground class, blue crosshairs show errors in building class, green crosshairs show errors in vegetation class.

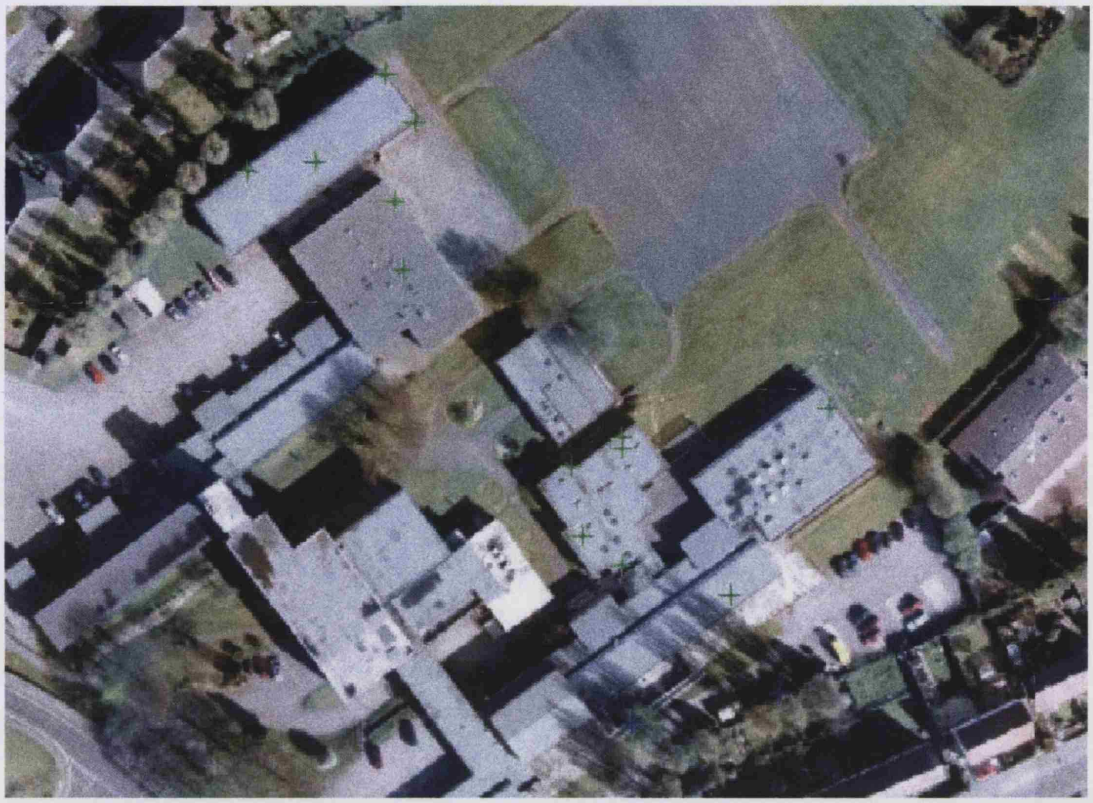

Figure 9.10: Zoomed in image of incorrectly classified vegetation points (green crosshairs) for the adapted geometric algorithm (linear 2m dataset)

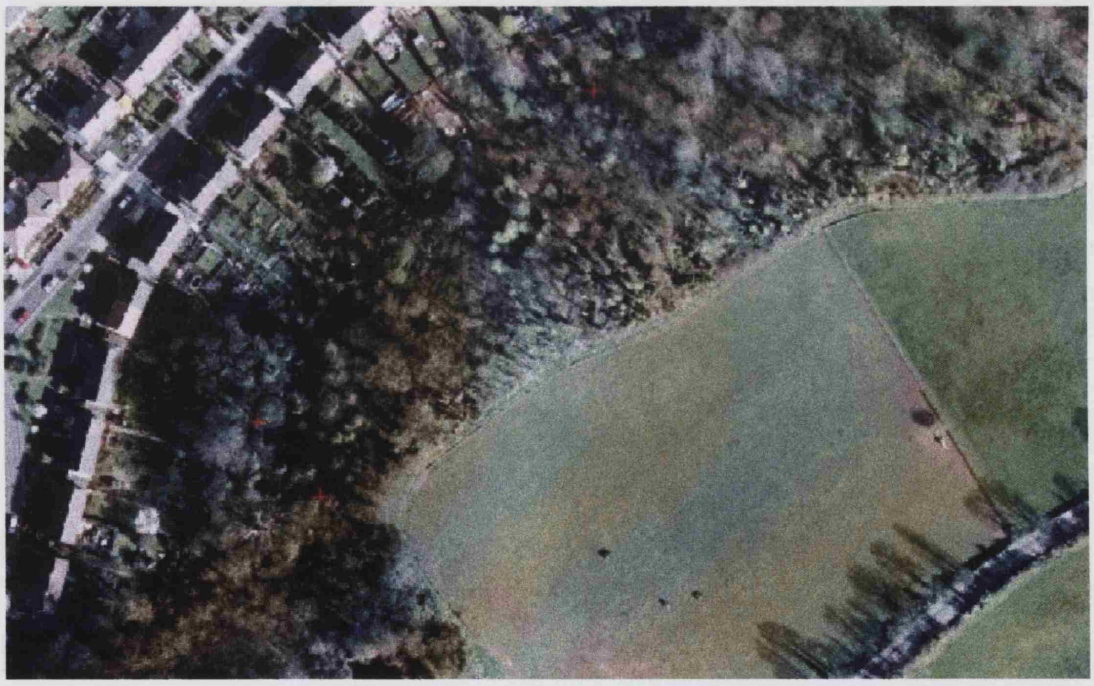

Figure 9.11: Zoomed in image of incorrectly classified ground points (red crosshairs) for the adapted geometric algorithm (linear 2m dataset)

Figure 9.12: Zoomed in image of incorrectly classified building points (blue crosshairs) for the adapted geometric algorithm (linear 2m dataset)

The observations above are qualitative. Further research could look at the correlation between a highly accurate classified aerial photograph and these clusters of errors. Such spatial analysis could also tell us whether the particular errors occur at the edges of certain features or land cover classes. In this approach, correlation between certain error types and the proximity to class boundaries could be investigated. Examination of this correlation was not performed as part of this investigation for the following reasons. First, any apparent patterning would be highly dependent upon the accuracy of the classification and this, in turn, is dependent upon the method used for classification. Assessing the relative accuracy of classifications is a research topic in its own right, with many authors promoting the use of particular types of classifier (e.g. Foody and Boyd 1999; Foody 2000; Tatem et al. 2001). A comparison of different methods therefore fell outside of the remit of this study. Secondly, even if a suitably accurate classification algorithm were available the only data available for the study area was a three-band (RGB) aerial photograph. Most successful classifiers require more than the RGB values alone. Given this, and the results for some trial classifications, it was considered that the classification accuracy was insufficient to sustain comparative investigation. However, this is an avenue for future research. If this were to be conducted a number of spatial analytical techniques could be employed in order to provide further information to users. These include investigating the clustering of points within certain classes, which could be performed using k-means analysis, or using spatial autocorrelation techniques such as Geary's C and Moran's I indices. The choice of which approach to use should also be related to user needs. The above solutions are inductive, data led methods. There is also enormous potential for the use of deductive techniques that assert the importance of the known spatial structuring of different feature classes (e.g. Barr and Barnsley 1997; 1998a; 1998b; 2000).

Potential future research strands include how this correlation information might be most effectively derived and used for decision making, how the use of different classifiers may affect the correlation results, and finally how the error patterns themselves might be used in classification algorithms as an additional source of information. If

it could be shown quantitatively that the magnitudes or types of errors were related to features on the ground (e.g building edges) this information could be used in classification. Future work should assess whether this information can significantly improve classifications.

9.4.3 Accuracy of Height Modelling

In order to assess how well the segmentation datasets can be used for the modelling of height, the classified points or pixels would ideally be used to compile discrete vector models of above ground features. However, this task is a complex one, and has been the focus of many previous investigations (e.g. Weidner 1995). The reconstruction of features from segmented points is a research topic in its own right, and developing methods for assessing the accuracy of vectors and reconstruction methods fell beyond the remit of this investigation.

In the absence of discrete, vector, models of the heights of the 'above ground' objects, in order to assess the accuracy of the object heights the only alternative was to use the heights as modelled in a normalised Digital Surface Model (nDSM) and to compare these with a reference dataset. An nDSM is a model of the height above ground (bare earth), and so whilst it is a continuous model, it does also provide heights of the above ground objects. These heights may be measured in the surface model and compared with a reference dataset and the differences calculated. An example of one of the nDSMs created in this investigation is shown in Figure 9.13.

A further constraint on this investigation was that it was necessary to model the segmented points as a continuous, grid surface, in order that they might be subtracted from each other. Therefore, there were some restrictions on the datasets and algorithms which could be compared in this way. Clearly the raw data could not be gridded without the consequent introduction of interpolation errors (as was shown in Chapter 6). In addition, the investigation was only possible for those algorithms which generated segmentation results as a grid, i.e the results produced by

Figure 9.13: Example normalised Digital Surface Model for the nearest neighbour 1m dataset created using the default IRI method. The image shows the relative changes in heights above ground surface: the lighter the pixel the higher it is above ground.

the geometric and IRI algorithms. Whilst the IRI algorithm does not differentiate between the above ground classes, it does nevertheless identify those points which are above ground. These results, plus the classes of points produced by the geometric algorithm, were both compared with a reference dataset. Given that not all of the algorithms and datasets could be compared, only a selection of the datasets was investigated in order to provide an indication of the general patterns of error.

The reference dataset chosen was a high resolution terrestrial laser scanning dataset, which was described earlier in Chapter 4. This was used to create a sample dataset of known heights of distinguishable objects. There were some constraints on the choice of points for the sample. Sample points had to occur on above ground objects, and be on surfaces which would have been visible from both the ground and the air. Examples of suitable points included those on the edges of building roofs. Clearly, there would be difficulties created by choosing points on vegetation, firstly because there was some time delay between the capture of the two datasets, and secondly the heights of vegetation features are constantly changing (either because of natural growth, or simply movement in the wind). To avoid the problems with vegetation heights, only those observations occurring on buildings were selected. Points on both the flat roofed buildings and the peaks of some pitched roofs were chosen. There were also constraints on determining the sample size. For any test to be reliable, a sample has to be representative of the population from which it is drawn. In the case of this investigation the size of the population (i.e. the number of points classified as 'above ground') was unknown. Given this, and the restrictions on the points which could be chosen, a manual sampling strategy was employed, and the number of sample points determined by time practicalities. As the sample size could not be representative of the population, and the fact that not all of the datasets could be compared, a small sample of 20 building points was created. Figure 9.14 shows how the measurements of the object heights were collected from the terrestrial laser dataset.

Each of the chosen nDSMs were compared to this reference dataset, the residuals

Figure 9.14: Showing the measurement of object heights within the reference dataset: example 1

were computed, and descriptive statistics recorded. No t-test was conducted to compare these results because the sample size was too small. The results are presented in Table 9.6. The datasets used in this analysis were the 1m resolution bilinear, bicubic, and nearest neighbour datasets for the adapted geometric algorithm and the default IRI algorithm. These were chosen on the basis that they contained the least error (according to the results presented in Chapter 6).

Table 9.6: Descriptive Statistics for Comparison of nDSMs with reference data

Dataset		Mean Error (m)	RMSE (m)
Geometric adapted			
	linear1m	0.71	0.92
	nearest 1m	0.40	0.70
	cubic1m	0.69	0.93
IRI default			
	linear1m	1.20	1.39
	nearest 1m	0.65	0.77
	cubic1m	0.80	1.13

From the results in Table 9.6 it would appear that the lowest height modelling errors for each of the algorithms were produced using the nearest neighbour datasets, with RMSE values being almost half that of the bilinear and bicubic counterparts for the IRI algorithm and at least 20cm lower than the others for the geometric algorithm.

The geometric algorithm was found to introduce lower height errors than the IRI method, this was slightly surprising given that the classification accuracies of the geometric method had been lower than the IRI method.

9.5 Answering the Research Questions

As with the previous chapter, the results from these investigations contribute to a broader analysis of a range of research questions. The answers to these broader research questions are likely to be of direct relevance to users of ALS data wishing to know which factors in the segmentation process can affect the sources and operation of errors which are introduced.

In order to compare the classification accuracies of the segmented datasets the z-test was used as the samples were again independent. Here the use of the z-test differs slightly to its application in the previous chapter. Whereas in Chapter 8 the test was used to compare the overall accuracy of classifications, here it is used to compare sub-components of each contingency table - and as such compare the producer's accuracy only. In this way it is possible to assess whether there are significant differences between the classification accuracies of the different segmentation methods. Differences between several datasets revealed answers to the following research questions.

9.5.1 Q1 Does using different segmentation algorithms cause a significant difference in accuracy?

The producer's accuracy results were compared for each of the classes in each of the datasets in order to assess whether there were significant differences between the segmentation methods. The full results are presented in Appendix J. The results indicated that in the vast majority of comparisons there was no significant difference

between the classification accuracies of the geometric and the ATM algorithms. Where significant differences did occur these tended to be in the comparisons of ground classes. The absence of significant differences between the classifications would suggest that for a user wishing to know which segmentation method to use for his/her application the choice of algorithm makes little difference.

This absence of significant differences was not anticipated - largely because the assumption that gridded data will produce lower quality results than raw, discrete data. Clearly, in terms of segmentation classification accuracy this was not, in fact, the case.

9.5.2 Q2 Does changing the parameter settings for the segmentation algorithms cause a significant difference in accuracy?

For this experiment the results from the two parameter settings for each of the algorithms were directly compared using the z-test. Results tables detailing the calculated values of z and the associated hypothesis results are presented in Appendix N. These results also show that there is generally an absence of any significant differences between the classification accuracies, indicating that the success of both algorithms is largely unaffected by changing the parameters/ settings of the methods.

9.5.3 Q3 Does the use of raw or gridded data cause a significant difference in segmentation accuracy?

This test was only applicable for the ATM method as this was the only method which used both irregular (raw) and gridded data. The producer's accuracy for the segmented raw dataset was compared with predictions from the other gridded datasets and statistical analysis of differences was carried out. The results of the z-test comparisons and resultant hypotheses results are shown in Tables 9.7 and 9.8 in

which the dash (-) symbol represents no significant difference, whilst a star (*) represents a significant difference. The results clearly show that in the vast majority of comparisons there was no significant difference in the classification accuracies of the raw and gridded datasets. This implies that there is no advantage to restructuring the raw data in terms of segmentation classification accuracy.

Table 9.7: Results of comparison of raw data segmentation with gridded data segmentation for the ATM method with default settings

		Class	
z-test Results	Dataset	Buildings	Vegetation
	Linear 0.5m	0.69	1.49
	Linear 1m	0.00	0.71
	Linear 2m	0.31	0.00
	Cubic 0.5m	0.74	0.13
	Cubic 1m	0.30	0.52
	Cubic 2m	1.69	0.65
	NN 0.5m	0.31	1.13
	NN 1m	0.98	0.89
	NN 2m	1.09	1.13
Hypotheses results			
	Linear 0.5m	-	-
	Linear 1m	-	-
	Linear 2m	-	-
	Cubic 0.5m	-	-
	Cubic 1m	-	-
	Cubic 2m	-	-
	NN 0.5m	-	-
	NN 1m	-	-
	NN 2m	-	-

9.5.4 Q4 Does the use of different data restructuring methods cause a significant difference in segmentation accuracy?

For this investigation it was necessary to compare the producer's accuracy results (for buildings and vegetation classes) for each of the gridding methods with each other, for each algorithm (e.g. linear 0.5m vs cubic 0.5m) for all possible combinations. The results of the z-test indicated that again for the majority of comparisons there was no significant difference between the accuracies of the gridded datasets. No differences occurred between the linear and cubic datasets; the only statistically significant differences found were between these and the nearest neighbour datasets,

Table 9.8: Results of comparison of raw data segmentation with gridded data segmentation for the ATM method with adapted settings

z-test Results		Dataset	Class Buildings	Vegetation
	Dataset	Linear 0.5m	1.28	1.05
		Linear 1m	0.00	0.48
		Linear 2m	0.66	0.00
		Cubic 0.5m	2.95	2.21
		Cubic 1m	0.33	1.85
		Cubic 2m	0.87	0.36
		NN 0.5m	0.22	1.60
		NN 1m	0.55	2.02
		NN 2m	0.76	1.05
Hypotheses Results		**Class**	Buildings	Vegetation
	Dataset	Linear 0.5m	-	-
		Linear 1m	-	-
		Linear 2m	-	-
		Cubic 0.5m	*	*
		Cubic 1m	-	-
		Cubic 2m	-	-
		NN 0.5m	-	-
		NN 1m	-	*
		NN 2m	-	-

which was caused by the lower 'above-ground' classification accuracy offered by the nearest neighbour dataset . Full results are presented in Appendix O. In practical terms these results suggest that users could save processing time and data volumes by using a bilinear interpolation algorithm to grid data, rather than a bicubic, in the knowledge that this will not significantly decrease the quality of their data.

9.5.5 Q5 Does decreasing the spatial resolution of the dataset cause significant differences in segmentation accuracy?

To answer this research question the producer's accuracy component of each of the contingency tables for each of the interpolation methods were compared. In other words, the contingency table for the segmented dataset produced using the linear 2m data for the geometric algorithm was compared with its 1m and 0.5m counterparts. Such comparisons indicated the existence of any significant differences between the classifications. These differences could only be caused by differences in

spatial resolution given that all other variables were equal.

The results of the z-test comparison revealed that there were few instances of significant difference in classification accuracy for the nearest neighbour datasets, and a general absence of significant difference in any of the methods in the comparison of 0.5m and 1m resolution data. The latter was thought to simply be the result of the similarity of the 1m and 0.5m datasets. Some differences will occur because the point spacing was $1.2m^2$, and as such the 0.5m dataset was slightly more than double the resolution. These differences were found to be very minor.

The results of this investigation are important to users wishing to decrease processing time and data volumes without decreasing the quality of the laser data.

9.5.6 Q6 Are there significant differences between the classification accuracies for each of the classes within the datasets?

This was an important investigation as the results may help users to determine which segmentation algorithm is the most appropriate to their particular needs. From the results of the comparison, which is shown in full in Appendix P, it was found that there are significant differences between the classification accuracies for the two 'above ground' classes. Table 9.4 shows these differences are largely because of the higher accuracies produced for the building class. The fact that there are significant differences between the classes indicates that different segmentation algorithms may be more suitable for some applications than others. It is the responsibility of the user to use information, such as that shown in Table 9.4, to determine which algorithm will produce the highest accuracies, and by implication introduce the least error during the segmentation phase of post-processing.

9.6 Summary of Findings and Discussion of Wider Relevance

Contrary to the conclusions from the filtering and restructuring investigations, it was found that, despite some qualitative differences, there were no significant differences in segmentation classification accuracies caused by using either different algorithms, different spatial resolutions, different input data, or different parameter settings. In practical terms this means that user choice of segmentation procedure is far less important in terms of the introduction of error, than either the filtering or the restructuring process.

In the comparison of the different patterns of error, it was found that there were some clear differences in the types of errors and their land use classes. Further research could determine if there were correlations between types of error and the classes of objects on the ground. Differences between the success of classifying different land cover categories were noted, indicating whether users should bear this in mind when selecting the segmentation algorithm most appropriate to their needs.

9.7 Chapter Summary

A comparison of segmentation algorithms and parameter settings is offered to determine which (combination of) methods introduces the least error into the resultant dataset. Errors are assessed in terms of their magnitude and spatial patterning for both classification and height modelling. The importance of understanding these errors is shown with reference to the application of the segmentation results for 3D reconstruction of objects.

Chapter 10 investigates how the results from the three investigations presented in Chapters 6 to 9 can be linked together to provide practical information to users of ALS data. The discussion in Chapter 10 culminates with an example of how this

information may be visualised and subsequently used in order to reduce or manage error.

Chapter 10

Discussion and Visualisation of Results

10.1 Chapter Overview and Aims

This chapter summarises and provides further context for the principal findings from the empirical research chapters (Chapters 5 to 9). The contributions to knowledge arising from these findings are discussed. The chapter also explores how the findings may be employed by ALS data users in three practical examples. Finally, a brief discussion about the use of visualisation techniques to improve our understanding of error and uncertainty is offered.

10.2 Summary of Principal Research Findings

The post-processing errors were calculated separately for three processes: restructuring, filtering and segmentation. The main findings regarding the characteristics of these errors, and how they relate to the expected range of errors are discussed below.

10.2.1 Restructuring errors

In the restructuring investigation it was found that high magnitudes of error were introduced regardless of which interpolation method or grid resolution were employed. Results from the pilot study calculated RMSEs of residuals ranging between 1.65 and 3.70 metres (based on the results from the 95% split-sample method for error estimation). For the main study the range of RMSEs for the residuals ranged between 1.25 and 13.69 metres.

Two observations may be made here. Firstly, the amounts of error introduced during this stage of post-processing are considerable and may severely affect the quality of any analysis based upon the surfaces created. Secondly, the large range of values suggests that using different interpolation methods or grid spacing can significantly alter the amount and spatial patterning of error within the resultant surface. Results suggested that the nearest neighbour algorithm produced the greatest overall error, and that the bilinear method produced the least. It was considered that neither the biharmonic spline nor the ordinary block kriging were suitable methods for the modelling of urban elevation surfaces. For the spline this was because of the problems of modelling the surface in unconstrained regions (such as at the edges of the dataset), whilst for the kriging it was suggested that it may not be a suitable method for modelling urban structures given that this method was initially developed for the modelling of naturally occurring phenomena. The grid spacing closest to the original point spacing of the ALS data produced the lowest errors.

Finally, the spatial patterning of errors suggested that the occurrence of the highest magnitude errors was correlated with the incidence of surface discontinuities. This was not surprising given that the inherent difficulty of interpolation in areas of discontinuities. Some of the differences noted in the characteristics of the errors over different types of discontinuity were nevertheless interesting, particularly with regard to the observation that the highest errors appeared to occur over vegetated areas and that these appeared to be more clustered than those errors associated

with building edge discontinuities.

For users, these findings would imply that the choice of interpolation method and grid spacing can both significantly alter the amount of errors incorporated within the resultant surfaces, and may also affect the distribution of the errors across the surface.

10.2.2 Filtering Errors

The results of the filtering investigation showed that the choice of algorithms, parameters, and input data structure can all significantly affect the amount of aggregate error within the resultant dataset. In terms of overall accuracy the geometric algorithm produced the highest classification accuracies, particularly when using data restructured using the nearest neighbour algorithm. The fact that the classification accuracies for the raw data were lower than those produced by some forms of interpolated data refutes the claim that restructuring data will always introduce more error into the resultant processed data. This is an important finding for the ALS community. Coupled with the fact that the highest classification accuracies were produced using a grid-based filtering technique, it suggests that ALS data should be restructured in order to improve filtering classification accuracies. Filtering should be carried out using a nearest neighbour interpolation routine. The relative amounts, and the spatial distributions of both Type I and Type II errors, differed substantially for the three filtering algorithms investigated. This suggests that, both the amount and the type of classification errors may be significantly affected by the choice of filtering method.

In terms of assessing the errors in the height modelling of the results of filtering, it was found that there was no improvement in using the original laser data, and that in many cases the use of interpolated data also improved the height accuracies of the models. The RMSEs of the estimated height errors ranged between approximately 3m for the geometric and IRI algorithms up to approximately 11m for the ATM

method. The magnitude of this difference indicates the importance of choosing the most appropriate filtering algorithm.

The spatial resolution of the input points was found not to significantly affect the amounts of error introduced. For ALS users this means that lower resolutions may be used for filtering without an associated increase in error. However, there will be some numerical limits on the effect of spatial resolution, it is suggested here that these are likely to be within $\pm$ 2 times smaller or larger than the nominal raw data spacing.

10.2.3 Segmentation Errors

There were no significant differences in the classification accuracies caused by using different segmentation algorithms, spatial resolutions, input data or parameter settings. For ALS data users this implies that the choices for the segmentation of laser data are much less important than those for either filtering or restructuring. There were, however, thought to be differences in the spatial patterning of different types of errors, the occurrence of which may affect a user's choice.

10.3 How could this information be used?

The findings presented in this thesis may be used in two different ways. Firstly the results could simply be used to aid future user decision-making based on the answers to the research questions that are presented at the end of each of the research chapters. For example, it may be sufficient for a user to know that there is no significant difference in errors caused by using gridded rather than the original configuration of data points, and using this information the user may define the data structure and algorithms they wish to use.

Alternatively, a user may wish to adopt the methodology presented here to identify

the characteristics of errors in a different study area. In this way the methodology could be used as a standard framework for all users to assess the errors incorporated during post-processing.

It is likely that most users would fall into the first category, and would be interested solely in the high-level findings from this research. This assumption is based on research such as that of Wechsler (2003) and Harrower (2003) who both analysed how and to what extent DEM users employ uncertainty information. Wechsler (2003) conducted a survey to investigate how DEM users address uncertainty. She found that 55% of respondents to a questionnaire considered that uncertainty was either 'very' or 'somewhat' important in their work, however of these respondents most reported that they would only be prepared to spend a minimal amount of time assessing DEM uncertainty in their investigations. These findings, coupled with the fact that there are relatively few tools for managing uncertainty in many commercial software packages, support the assumption made here that most users will only use the very high level summary results from this research.

The high level results (summarised in Section 10.2) suggest that different algorithms/input data/resolution can create significantly different magnitudes of errors, and may produce different patterns of errors. In other words, the laser data user may make certain choices regarding the algorithms and data structures that he/she employs during post-processing in order to minimise the errors which are introduced. The choice of which algorithm or data structure or spatial resolution to use depends on the output model required as well as the application for the model. The choices that the user has to make can be visualised rather like a tree, through which the user must determine a pathway. This is illustrated in Figure 10.1. The pathway through the choices is likely to be largely defined by the application for which the user is processing the laser data. Some example 'pathways' for different applications are discussed here.

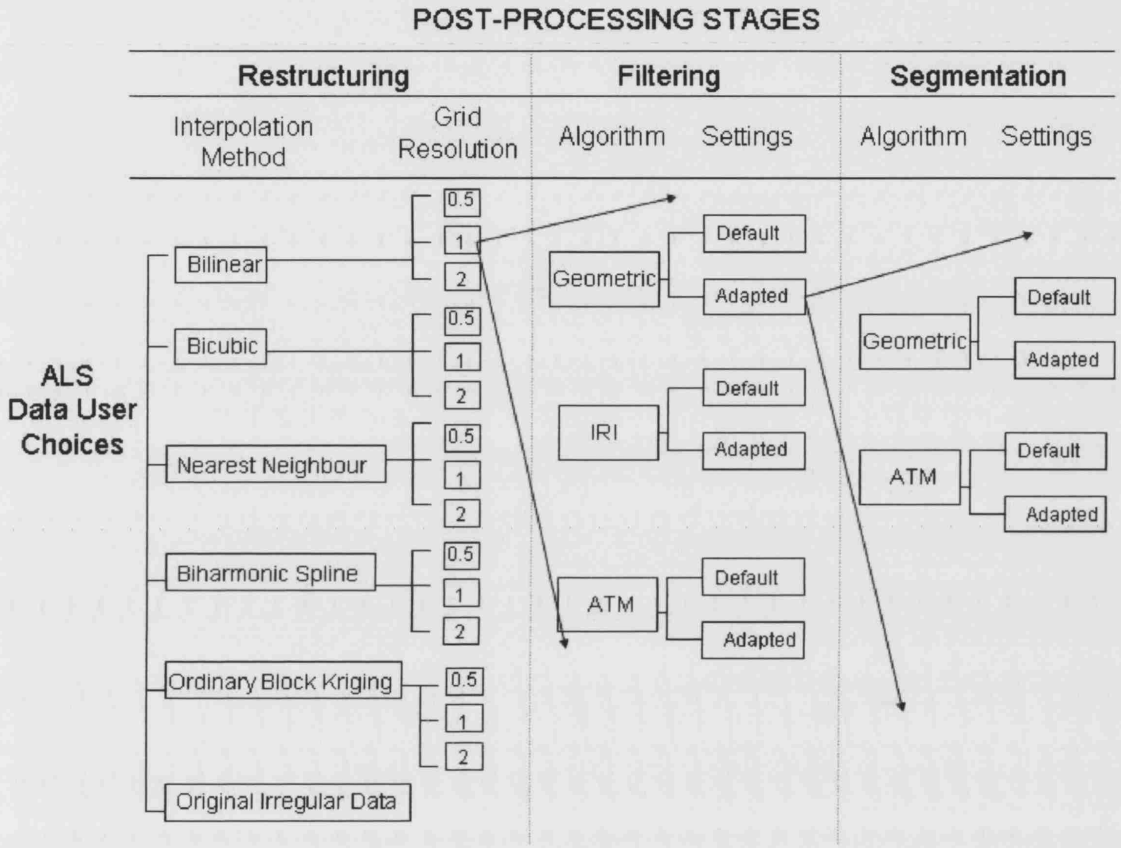

Figure 10.1: Showing the range of choices the user must make during post-processing of ALS data. For each of the boxes resulting from the restructuring stage, there are the full range of options indicated for filtering. These filtering choices are shown only once (for simplicity) for one of the boxes. The same is true for the segmentation choices shown, with each relating the results from each of the filtering options.

10.3.1 Example 1

On the basis of the results presented in this thesis, there are three suggested pathways for a user who wishes to create three separate models (a DSM, a DTM, and a segmented dataset for use in 3D reconstruction). These are shown in Figures 10.2 to 10.4, and are explained further here.

For the user who wishes to solely create a DSM, a bilinear interpolation method at a grid spacing which is close to the laser point density should be used (as shown in Figure 10.2). However, this methodology is not appropriate for the subsequent creation of the DTM. In order to minimise the errors introduced the user should reprocess the data, this time using a nearest neighbour interpolator to restructure the data at 2m resolution, and then filter these data using the adapted geometric algorithm. This pathway is depicted in Figure 10.3. Finally, for the creation of a segmented dataset, the user should use the results from the filtering stage and employ the adapted ATM algorithm for segmentation. Whilst the results from the ATM algorithm were not found to be significantly better than those in the geometric algorithm, the results from this investigation did show some marginal improvements (Figure 10.4). So, in this example, in order to reduce error in the creation of these three datasets/models, the user should process the ALS data twice using different pathways through the post-processing flowline in order to reduce the amount of error introduced.

10.3.2 Example 2

Users may also be interested in the spatial patterning of errors which was investigated within this research. The findings may influence their choice of path through the post-processing. For example, a user who may want to create a DSM but is interested principally in the aesthetics of the surface may choose either a spline or a geostatistical interpolator to create the surface. Results from this investiga-

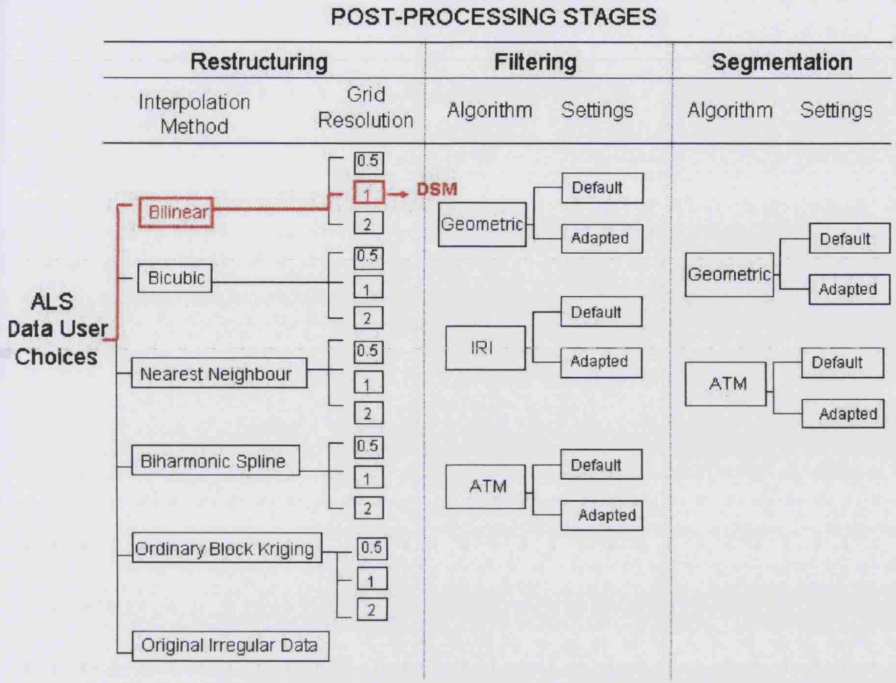

Figure 10.2: Showing recommended path for user wishing to create a DSM incorporating the least overall error

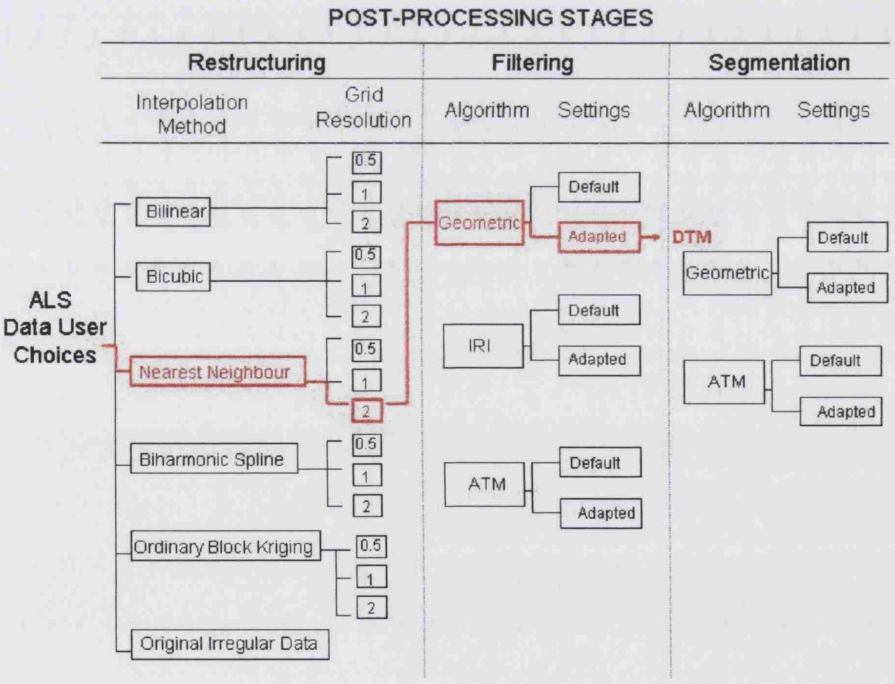

Figure 10.3: Showing recommended path for user wishing to create a DTM incorporating the least overall error

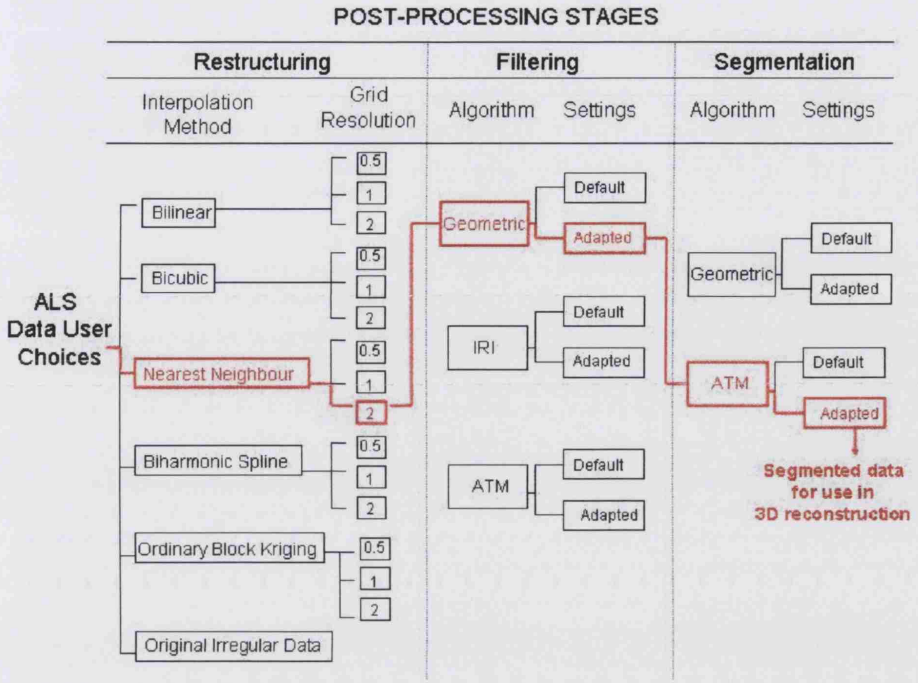

Figure 10.4: Showing recommended path for user wishing to create a segmented dataset which incorporates the least overall error

tion showed that the biharmonic spline interpolator produced spikes at the edges of the dataset, and given this the ALS data user may decide that a geostatistical interpolator may be more appropriate.

10.3.3 Example 3

Users may also be less interested in minimising overall error and instead be more concerned with reducing errors in selected areas during data processing. For these users, information regarding the relative success of the individual classifications from segmentation algorithms will be of interest. Users may also be interested in the different types of errors introduced during processing, including errors in classification and height modelling for both the filtering and segmentation stages. Depending on which type of errors are more important to a given user, there will be different recommended paths through the post-processing. For example, for a user more interested in the accuracy of height modelling of above ground features, the

geometric algorithm should be used with a dataset restructured using the nearest neighbour algorithm at a 1m spatial resolution. For a user to whom the accuracy of the classification of above ground points is more important then the ATM algorithm should be used for the processing.

10.3.4 Discussion

These examples demonstrate that there is no single optimal pathway through the choices for ALS post-processing. The pathway most suitable for a particular user depends on the application for which the processing is designed. On the basis of the results presented in this thesis, different pathways will be recommended for those users interested in minimising overall error, different types of error, and the location of error. There are, of course, other factors which may influence a user's decision path. These include processing time for the algorithms, data volumes, and the varying costs associated with each of these. All of these factors, including the results from this study, could be used as a decision-support tool to aid ALS data users in the design of their post-processing pathway. This tool could be as simple as asking the user to define requirements such as which factor (error, speed of processing, data volume) is the most important. If the amount of error is of most importance to the user then the results from this investigation could be used, in combination with additional information regarding the user requirements, to suggest the most appropriate post-processing workflow.

This section has discussed how the majority of users may use the results from this investigation. However an additional type of user has been identified for this research. These users may use the methodology presented here to estimate the errors for their own study areas. For these users there is an additional issue which requires some discussion here, namely the communication of the errors and associated uncertainty.

10.4 How can this error information be communicated?

The fact that different algorithms, resolutions, and data structures can produce different patterns of errors indicates that there is some uncertainty regarding the modelling of the urban elevation surface across the study area. MacEachren (1995) and Ehlschlaeger and Goodchild (1994) suggest that there are a number of ways in which uncertainty may be communicated to the user. Popular methods include the use of animation, symbology, or hue, saturation and intensity of colours within a map (e.g. Hengl 2003). There are two schools of thought regarding the visualisation of uncertainty: some authors consider that uncertainty or error information should be integrated within the map of the phenomenon of interest, using what are termed 'integrated symbols' (e.g. Leitner and Buttenfield 2000; Edwards and Nelson 2001), whilst others consider that these symbols serve only to clutter the map (McGranaghan 1993).

Foody and Atkinson (2002) proposed that for remote sensing the spatial output of a process should be (at least) twofold: (i) a map of the variable of interest and (ii) some assessment of the uncertainty in that map. Following this recommendation, an attempt was made to visualise the uncertainty in the surfaces created in the restructuring investigation. To do this, the values of the calculated residuals (which were estimations of the error) were used to interpolate another surface designed to represent the spatial uncertainty within the surface models. Whilst the exact form of this resultant surface will alter depending on the method used for interpolation, the example shown here demonstrates how the surface error can be used to communicate generalised uncertainty information to users. Figure 10.5 shows the bilinear surface created in the pilot study, and Figure 10.6 displays the 'uncertainty surface' (created from the residuals) associated with this. Figure 10.7 shows a similar uncertainty map, for that associated with a bicubic surface. There are some similarities between the two uncertainty surfaces: in particular they both show that the principal areas of uncertainty occur in the region of vegetation. However, it is the differences in the patterns which can reveal information about the potential problems within the

underlying datasets. For example the uncertainty within the bicubic data is more concentrated than those in the bilinear, perhaps suggesting that there may be some spikes within this surface.

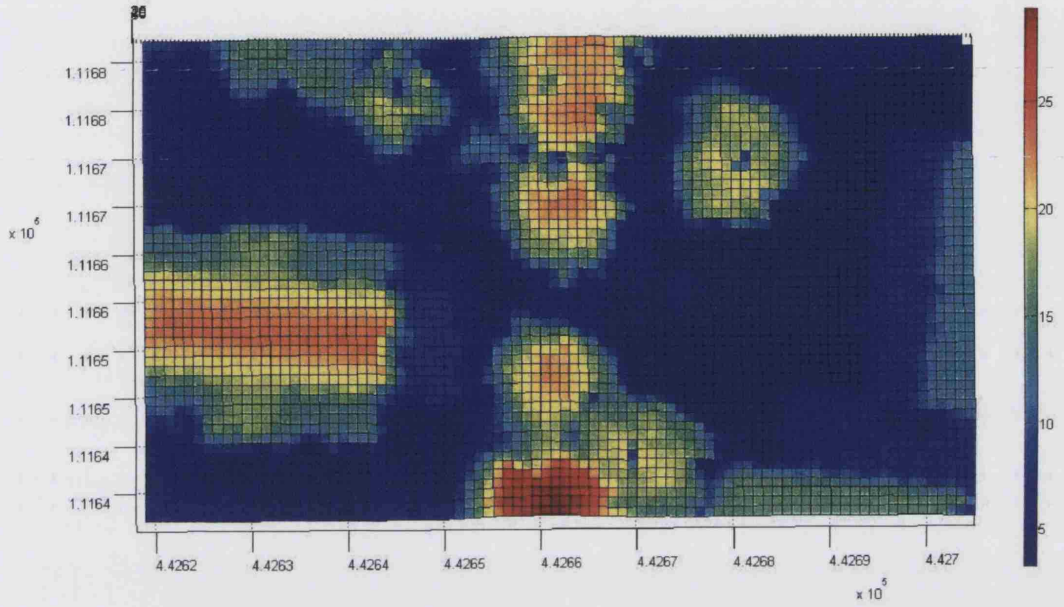

Figure 10.5: Plan view of surface created using linear interpolation. X and Y coordinates in metres

The surface for the bilinear 1m data in the main investigation was also displayed in this way (Figure 10.8), although the error surface is less clear here because of the much larger size of the study area. Figure 10.9 shows the enlarged NW quadrant of this residual surface: the type of error is shown by the colour of the filled contours and it can be seen that there are both positive and negative errors. Only the magnitude of the error can be used as an indication of uncertainty.

In this way, these uncertainty surfaces themselves could be used with relative ease to provide a rapid visual indication of the likely degree of success of any particular algorithm for any of the post-processing processes investigated within this research. What is proposed here is that without error and uncertainty information users cannot currently make informed decisions regarding their choice of post-processing methodology. Further work is recommended to investigate how this information can

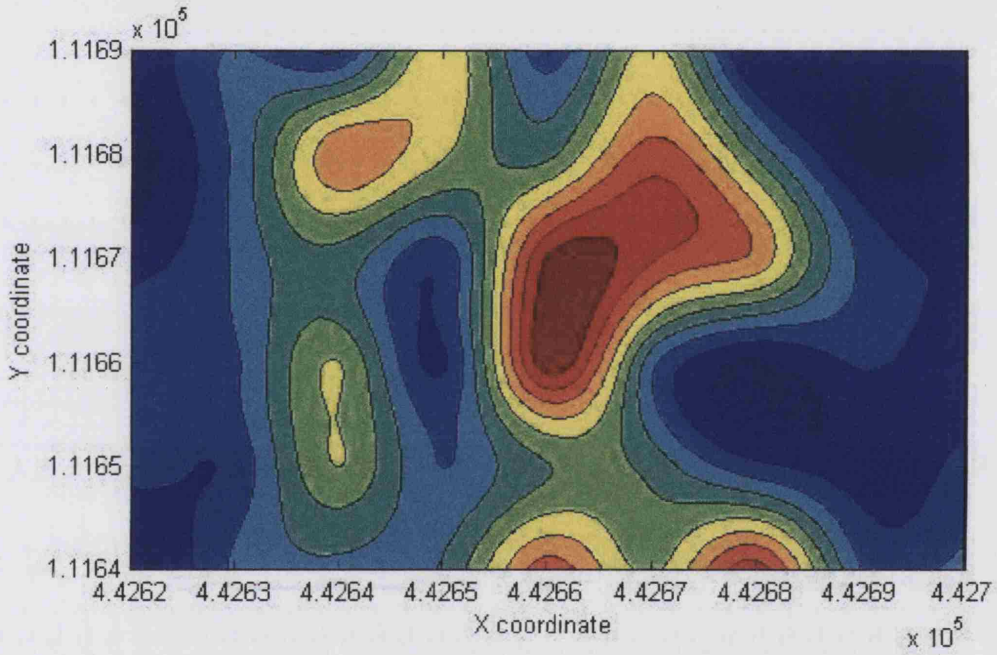

Figure 10.6: Plan view of 'uncertainty' surface for the DSM created using bilinear interpolation (1m dataset). For legend see figure 10.7

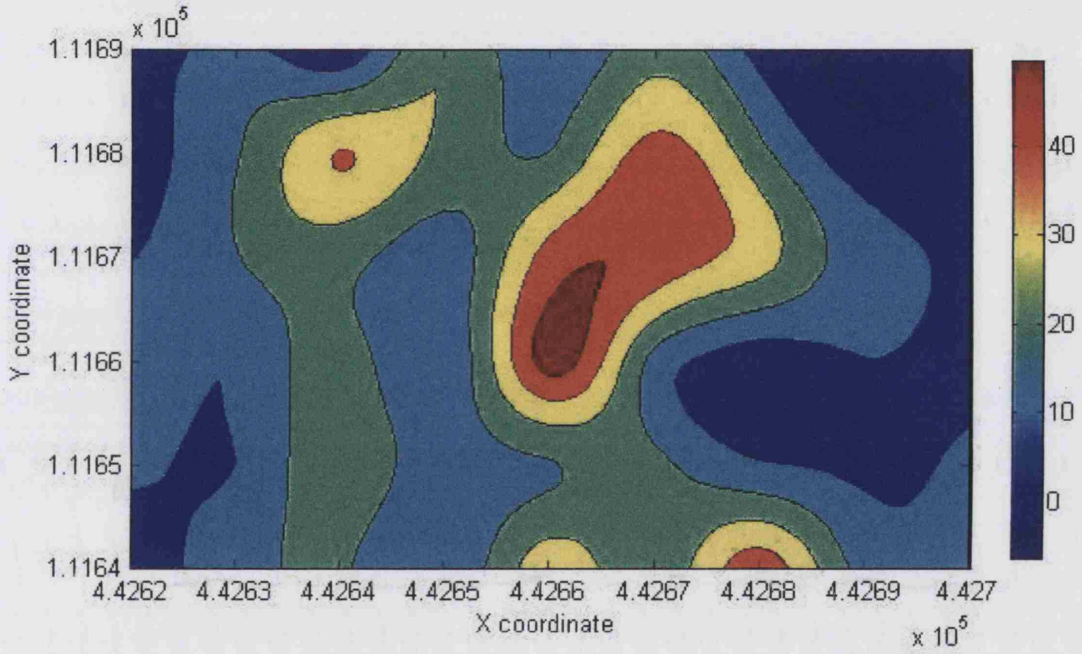

Figure 10.7: Plan view of 'uncertainty' surface for the DSM created using bicubic interpolation (1m dataset). Note that legend denotes magnitudes of errors (m) multiplied by 10 for display

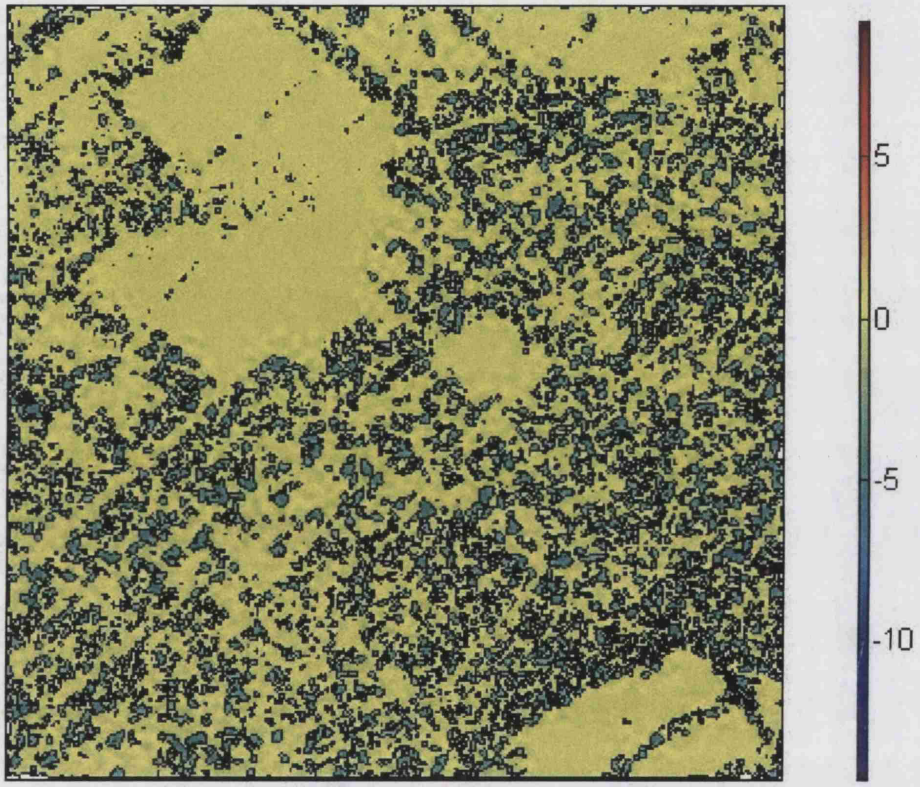

Figure 10.8: Plan view of uncertainty surface for the DSM created using bilinear interpolation (1m dataset) for the main study area. Amounts of probable error displayed are in metres

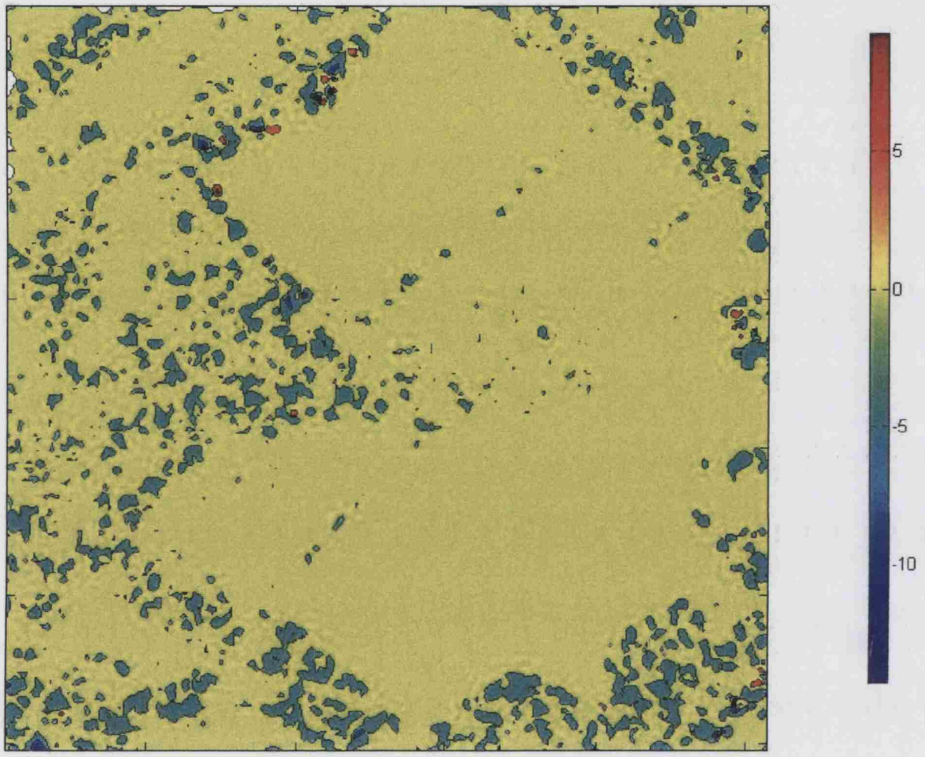

Figure 10.9: Plan view of the NW quadrant of the uncertainty surface for the DSM created using bilinear interpolation (1m dataset) for the main study area. Amounts of probable error displayed are in metres

be incorporated within the user decision making process, and to assess the relative merits of different representation methods.

10.5 Chapter Summary

This chapter has summarised the main technical findings in this thesis, and has shown how these may be used to aid a user's choice of algorithms and parameter settings for each of the post-processes investigated in order to reduce error. The following chapter explores areas for further work which have been touched upon in this, and in previous chapters.

Chapter 11

Conclusions

> Data may pass through many different transactions and custodians, each providing their own protocols or interpretations....uncertainty is not a property of the data contents so much as a function of the relationships between data and users. *Zhang and Goodchild* (2002) page 5

11.1 Contribution to the Research Literature

This thesis has sought to develop a balanced appraisal of the sources and operation of errors in the processing of ALS data. The main intellectual driver for the research was the lack of in depth information regarding the characteristics of post-processing errors, which was seen to be in quite stark contrast to the much larger body of research centred upon understanding the supplier processing errors. In the investigation of the post-processing errors, this thesis has developed: a quantification of overall post-processing errors; the identification of different types of errors; and some analysis of the spatial patterns of the errors for each of the three main stages of ALS data post-processing. In this way, the results may be used to inform laser data users regarding which factors they should consider in order that they may reduce

error in their post-processing.

The thesis also presents a framework for measuring, and communicating, error which, it is suggested here, could be used as a standard framework for all users of ALS datasets. Greater standardisation of procedures for mapping and modelling ALS errors is required in order that results may be compared. Finally, a new filtering algorithm, the geometric method, is also proposed. This was found in some instances to introduce less error than commercially available algorithms.

The future of laser scanning processing is currently heavily weighted towards the processing of original laser points. However, this study has shown that in terms of filtering and segmentation, this may not be the optimal developmental path. It may, in fact, be beneficial to incorporate a data restructuring stage within DTM post-processing in order to reduce classification errors, even if the points are then processed as discrete points rather than as part of a continuous model.

11.2 Applicability of the Results

The results of this study will be of use to the following sets of users:

- Most importantly, for users who have purchased raw laser data and need to process their dataset in order to produce a model, the results from this investigation support the decision-making so that errors may be managed, and minimised
- For those users wishing to purchase pre-processed data from a supplier, this study indicates those areas in the post-processing chain where error may have been introduced. In this way, the user will be able to make a more informed estimate regarding the quality of any analysis based on the processed ALS data
- For users of surface models (whether these be models of height, populations, or climate) the methodology presented here provides a robust framework which

may be used to identify and quantify errors.

11.3 Recommendations

The findings from this study emphasise the need for the quality of datasets to be both quantified and communicated. This quality audit must necessarily monitor all stages of data processing in order that defensible analysis might be conducted. The thesis calls for the supply not only of global quality statistics, but also of local error information, and demonstrates how this may be employed by users in their decision making. It is strongly recommended that local error information, along with the spatial pattern of error, should be provided to users by data suppliers.

11.4 Further Research

To extend the ideas presented within the thesis, there are a number of areas where further research should be conducted. These are discussed below.

11.4.1 Error propagation

The most obvious extension of this work is to define the function which can statistically describe the propagation of errors within the post-processing workflow. In Chapter 2 a full error budget equation was proposed:

$$\varepsilon_{LP} = f(\varepsilon_{cal}, \varepsilon_{GPS}, \varepsilon_{INS}, \varepsilon_{mm}, \varepsilon_{s}, \varepsilon_{R}, \varepsilon_{T}, \varepsilon_{sm}, \varepsilon_{ds}, \varepsilon_{f}, \varepsilon_{s}, \varepsilon_{rcon}) \qquad (11.1)$$

The characteristics of the errors introduced at each stage of post-processing have been investigated within this research. However, the modelling of the propagation of

the error from one process to the next lay beyond the remit of the study. A definition of the form of the function (f), would further aid understanding of the behaviour of the errors. In this way the errors may be predicted, and so managed and potentially reduced. Such a statistical investigation should not concentrate solely on the post-processing errors, but should look at the error propagation throughout the whole flowline, from data capture to the reconstruction and modelling of objects.

11.4.2 Visualising and understanding the spatial patterns of errors

It was suggested in the results chapters that the location of higher magnitude errors, which are created during the post-processing of ALS data, was related to the occurrence of above ground features in the landscape. In this way, the introduction of errors within the dataset should not necessarily be seen as a bad thing, for the errors may provide information which could, for example, enhance the results of an image classification algorithm. One potential application of this information is described below.

Different error characteristics and patterns may apply to different land covers: for example those errors over vegetation may be of a higher magnitude and may be more clustered, whilst those over buildings may occur in a more linear pattern. Further research should investigate the patterning of these errors to assess whether there are characteristics which relate to the land cover. Errors could subsequently be classified based on their identified characteristics, and these could then be used as an additional layer of information in a classification algorithm. Any enhancements in the accuracy of the classification could be determined based on a comparison of the algorithm in which the error information was not used during classification.

11.4.3 Repeat testing

As with any experiment, it would be useful to run repeat tests using different study areas to cover a wider variety of terrain types. This investigation has focused on two urban areas, and it would be interesting to assess the transferability of the findings to other terrain types, built morphologies and street configurations.

It would also be a valuable extension to assess whether using higher density point datasets can affect the characteristics of the errors introduced. Whilst technical developments in the hardware field are producing increasingly high densities of points, this does not affect the significance of the findings from this investigation for two reasons. First, ALS has considerable potential for the mapping sector of the market. For the capture of large areas of data (such as at a national scale) the processing of very high point density datasets would be unrealistic. Secondly, many users will simply not require high point densities for the majority of applications (including the 3D reconstruction of objects). For these reasons it is envisaged that many data suppliers will continue to supply quite low density data (1-4 points per m^2). The results from this investigation will therefore continue to be of relevance for the foreseeable future.

11.4.4 Additional extensions

This study has focused on error for decision making, but could be expanded to include other factors such as processing time and data volumes, in order to assess the relative importance of these factors to ALS data users. Each of these factors could be modelled within a decision support system in order to aid user decision-making for post-processing workflow design more formally than was suggested in the previous chapter.

In a similar vein, this research would also benefit from some further investigation regarding how error information can be effectively communicated to the data user.

This would involve some cognition studies related to how users perceive and manage error. In line with this, it is recommended that data suppliers should, in the future, aim to produce spatial error information as well as the original phenomenon of interest.

11.5 Final Word

The rapid developments in laser scanning technology are leading to increasing data capture accuracies. The same emphasis *must* be placed on developing more accurate techniques for the processing of laser data in order that the full benefits of this data source be realised. Even with future improvements, error will never be eliminated completely from ALS data processing. It is hoped that if this thesis has demonstrated anything, it is that the errors created during post-processing cannot and should not be ignored. As such, there should be continued research in this area to provide more information about the causes and behaviours of post-processing errors.

Appendix A

List of Publications From Thesis

A.1 Published Articles

Smith, S.L; Holland, D.A; Longley, P.A (2005) Quantifying Interpolation Errors in Urban Airborne Laser Scanning Models. *Geographical Analysis Special Issue 'Advances in Interpolation'.* In Press

Smith, S.L; Holland, D.A. and Longley, P.A. (2004) The Importance of Understanding Error in Lidar Digital Elevation Models *International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, XXXV , page 996-1001

Smith, S.L; Holland, D.A; Longley P.A (2003a) Investigating the Spatial Structure of Error in Digital Surface Models. *International Archives of Photogrammetry and Remote Sensing* XXXIV, Part 3/W13. Proceedings of ISPRS 3-D Reconstruction from Laser Scanning and InSAR; Dresden, Germany, page 136 - 142

Smith, S.L; Holland, D.A; Longley P.A (2003b)The Effect of Changing Grid Size in the Creation of Laser Scanner Digital Surface Models) *Proceedings of Geocomputation 2003*, page 137-142

Smith, S.L; Holland, D.A; Longley P.A (2003c) Interpreting Interpolation: The pattern of interpolation errors in Digital Surface Models derived from Laser Scanner

Data. *Proceedings of the GIS Research UK 11th Annual Conference* page 137 - 142

Smith, S.L (2003) Urban Remote Sensing: the use of LiDAR in the creation of physical urban models. Chapter 9 in *The CASA Book of GIS* (ed. Paul Longley), page 171-190

A.2 Recently Submitted Articles

Smith, S.L; Holland, D.A; Longley, P.A (2005b) An Analysis of Errors Introduced During the Post-Processing of LiDAR Data. Submitted to ISPRS conference June 2005 3D Mapping from Lidar and InSAR; Banff, Canada

Smith, S.L; Holland, D.A; Longley, P.A (2005c) Analysing the Effect of Restructuring Airborne Laser Data on the Quality of Digital Terrain Models. Submitted to ISPRS conference September 2005 LaserScanning 2005; Enschede, The Netherlands

Appendix B

Code and Functions

B.1 Code for interpolation of laser data onto regular grids at different resolutions

```
function [jk_95,jk_5, jk,back,stats,s] = grid_lidar (I,llx, lly, urx,ury);


warning off MATLAB:griddata:DuplicateDataPoints

[jk_95, jk_5] = jackknife (I,0.95);
[Xt1, Yt1] = meshgrid (llx : 1 : urx, lly : 1 : ury);
[Xt2, Yt2] = meshgrid (llx : 2 : urx, lly : 2 : ury);
[Xt4, Yt4] = meshgrid (llx : 4 : urx, lly : 4 : ury);
interp_technique = {'linear', 'cubic', 'nearest', 'v4'} ;
back_technique = {'linear', 'cubic', 'nearest', 'spline'};
grid_space = {Xt1, Yt1; Xt2, Yt2; Xt4, Yt4};
for it = 1:4
    for gs = 1:3
    jk{it,gs} = griddata (I(jk_95,1), I(jk_95,2), I(jk_95,3),
grid_space {gs,1}, grid_space {gs,2}, interp_technique{it});
    back {it,gs} = interp2 (grid_space{gs,1}, grid_space{gs,2},
jk{it,gs}, s(jk_5,1), s(jk_5,2), back_technique{it});
    stats {it,gs} = statistics (I(jk_5,3), back {it,gs}(:,:));
    end
end
```

B.2 RMSE function

```
function r = rmse (x,y)
\% rmse: calculates the root mean square error between two datasets
\%
\% Use function to calculate a measure of the
\% difference between two datasets
\% usage: rmse (x,y)
\% where x and y are arrays of equal size
\% The following inputs are required:
\% slsmith@ordsvy.gov.uk
\%
if sum(size(x) == size(y)) ~= ndims(x) ,error(
'x and y arrays must be of equal length');end
x1 = x(:); y1 = y(:);
Tx = isnan (x1); Ty = isnan (y1);
f = find ((Tx == 0 ) & (Ty == 0));
clear Tx; clear Ty;
x = x1 (f,:);
y = y1 (f,:);

r = sum ((x - y).^2)/ length(x);
r = sqrt (r);



\%r = sum ((x - y).^2)/ prod (size(x));
```

B.3 Subset funtion

```
function [s] = subset (I, llx, lly, urx, ury);
\% subset creates a smaller subset of large data file
\% I is the large file from which the subset is to be made
\% I must consist of x,y,z


n = size(I,2);
if n <3, error('[I] must contain at least three columns');end

s3 = find ((I(:,1) >= llx) & (I(:,1) <= urx) & (I(:,2)
<= ury)&(I(:,2) >= lly));
s = I(s3, 1:3);
```

```
\section{Statistics function}
function [s] = statistics (original, interpolated)

if size (original,2) ~= 1 | size (interpolated,2) ~= 1, error
('x and y must be vectors');end
if length (original) ~= length (interpolated), error
('x and y vectors must be of equal length');end
f = find (~isnan(interpolated));
diff = original(f,1) - interpolated (f,1);
s (1,1) = max (diff);
s (2,1) = min (diff);
s (3,1) = mean (diff);
s (4,1) = std (diff);
s (5,1) = rmse (original (f,1),interpolated(f,1));
```

B.4 Cross validation routine

```
function [all_residuals,back, gd,s2] = bristolcv_1mlidar(I,llx,lly,urx,ury);
\% LOO function stands for Leave-One-Out, which splits a dataset into two,
\% whereby the first subset contains one value, and the larger subset
\% contains all other values. The process works iteratively until all
\% possible subsets (of 1) have been created.
\% This function is used in the CROSS VALIDATION routine to assess the
\% success of resampling routines
\% This function uses EXCLUDE
warning off MATLAB:griddata:DuplicateDataPoints
\% tic

\%num_pts = length (I);
num_pts = 1;

s2 = cell(num_pts,1);
[Xt1, Yt1] = meshgrid (llx : 1 : urx, lly : 1 : ury);
[Xt2, Yt2] = meshgrid (llx : 2 : urx, lly : 2 : ury);
[Xt4, Yt4] = meshgrid (llx : 4 : urx, lly : 4 : ury);
interp_technique = {'linear', 'cubic', 'nearest', 'v4'} ;
back_technique = {'linear', 'cubic', 'nearest', 'spline'};
grid_space = {Xt1, Yt1; Xt2, Yt2; Xt4, Yt4};

for i=1:num_pts
    s1(i,:) = I(i,:);
    \% next need to find all values in I which do not equal s1, and put
```

```
   \% these into a cell array
    s2{i,1} = exclude(I,i);
    N = s2{i,1};

    for it = 1:4
        for gs = 1:3

            gd {it,gs} = griddata (N(:,1), N(:,2), N(:,3),
grid_space {gs,1}, grid_space {gs,2}, interp_technique{it});
            back(i,it,gs) = interp2 (grid_space{gs,1},
grid_space{gs,2}, gd{it,gs}, s1(i,1), s1(i,2),
back_technique{it});
            residual = back(i,it,gs) - s1(i,3);
            all_residuals(i,it,gs) = residual;



        end
    end

end


residual_abs = abs(all_residuals);
residual_abs(isnan(residual_abs)) = 0.0001;


\% to account for zero values
v = 0.0001;
f = find(~residual_abs);
residual_abs(f) = repmat(v,size(f));

for it =1:4
    for gs = 1:3

            figure
            scatter (I(1:num_pts,1),I(1:num_pts,2),(residual_abs(:,it,gs))
*10,'filled'); hold on;
            xlabel('x-Coordinate'); ylabel ('y-Coordinate'); title
('Scatter plot of model errors');

    end
end

saveas (figure(1), 'Scatter.eps');
```

```
\% Plotting the actuals against the predicted values to analyse trend

for p = 1:12
    figure

        it = floor ((p+2)/3)
        gs = p - (it-1) * 3

        plot (I(1:num_pts,3),back(:,it,gs),'+'); hold on;
        xlabel('Actual value'); ylabel ('Predicted value'); title
\%('Plot of actual against predicted values of z');
       \% fit line of best fit
      \% create variable string
\% saveas (figure (p + 1), 'Figure.eps');

    end

end
 save

\%toc
```

B.5 Function used for randomly spliting dataset

This function was written by Dr Isabel Sargent, Research Innovation, Ordnance Survey

```
function [r,D]=jackknife(I,pc);
\%jackknife randomly split a dataset into two
\%
\% usage: [r,D]=jackknife(I,pc);
\%
\% I is the matrix of data, where each row is an observation and each column
\% is a variable
\% pc is the percentage of the observations to be excluded (between 0 and 1)
\%
\% r is a a vector of the row number selected from I. It represents
\% pc percent
\% of the data
\% D is the remaining (1-pc percent) of the data
\%
```

```
\% isargent@ordsvy.gov.uk

[obs,vars]=size(I);
numout=round(obs*pc);
ran=rand(obs,1);
[out,idx]=sort(ran);
r=idx(1:numout);
D=idx(numout+1:end,:);
```

B.6 Exclude funtion used for cross validation routine

Please note this funtion was not written by the author of this thesis

```
function res = exclude( data, inds )
\% EXCLUDE --- exclude data from a vector
\% exclude( data, inds )
\% leave out of <data> the points at the indicated indices
\% e.g., exclude(data, 5) to leave out the fifth point
\%       exclude( data, [2 10]} to leave out the second and tenth
\% if data is a matrix, this assumes that each row is one data point

\% (c) 1998-9 by Daniel T. Kaplan, All Rights Reserved

indices = 1:(length(data));

for k=1:length(inds)
   indices = indices(find(indices ~= inds(k)));
end

[r,c] = size(data);

if c == 1 | r == 1
   res = data(indices);
else
   res = data(indices,:);
end
```

Appendix C

Workflow for Matching Individual Scans

This appendix details the workflow used in the PolyWorks software package for aligning the individual terrestrial laser scans. This work was conducted by the data suppliers, 3D Laser Mapping.

C.1 Characteristics of the Raw Data

- Data captured in an urban environment itemGPS co-ordinates for the scanner position
- Overlapping scans
- The GPS co-ordinates are transformed into National Grid co-ordinates using Grid Inquest
- The 7 least significant digits of these are put into individual transformation matrix format used by PolyWorks

C.2 Workflow in Polyworks IMAlign

- The project file from RiScanPro is accessed and the scans imported

- The matrix for the scan position is applied to the scan
- All scans are translated down by the antenna offset
- The scans are constrained to rotate about the Z axis (the scanner will have been nearly level) to line them up to the right direction
- The rotation constraint changed to centre on the offset antenna position, as 3D constraint with the tolerance of the GPS quality
- Each adjoining scan is looked at in turn and a few common features are marked in each, these are used by PolyWorks to move the scans so that they match better
- A best fit alignment is the performed where the scans are rotated and translated within the given constraints to the best possible fit and locked
- After a few have been aligned in this way the position lock is removed and an overall best fit is done using the constraints on the group
- When this has been done on all the scans the alignment matrices are exported from IMAlign and imported into RiScan Pro for each scan position (SOP)
- The larger numbers of the National Grid co-ordinates are put into a project matrix (POP) itemThe data is exported in text format in the project co-ordinate system

Appendix D

Digital Surface Models

D.1 Digital Surfaces Models Created Using Different Restructuring Techniques

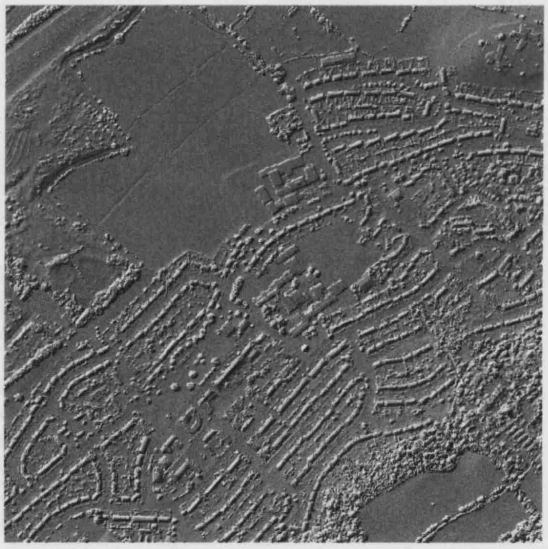

Figure D.1: Bilinear 0.5m DSM

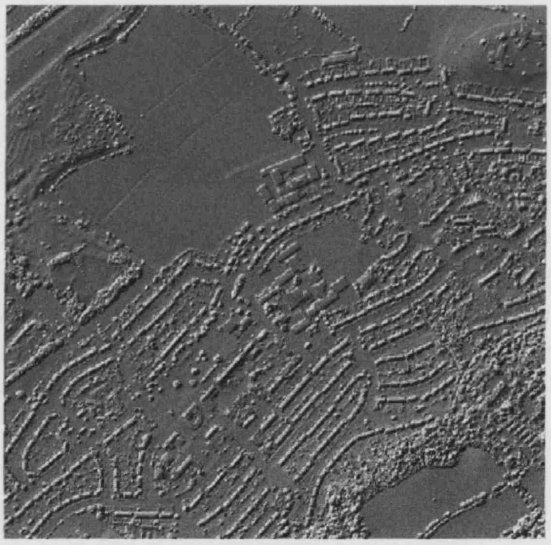

Figure D.2: Bilinear 1m DSM

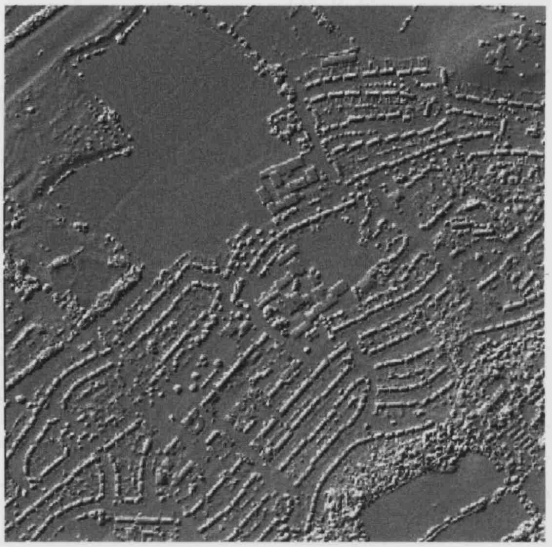

Figure D.3: Bilinear 2m DSM

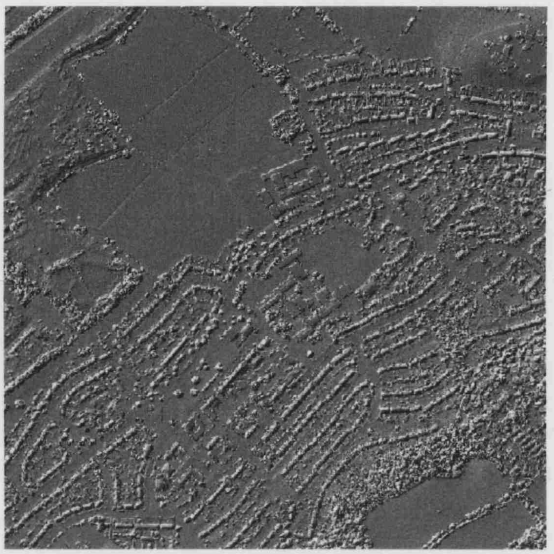

Figure D.4: Bicubic 0.5m DSM

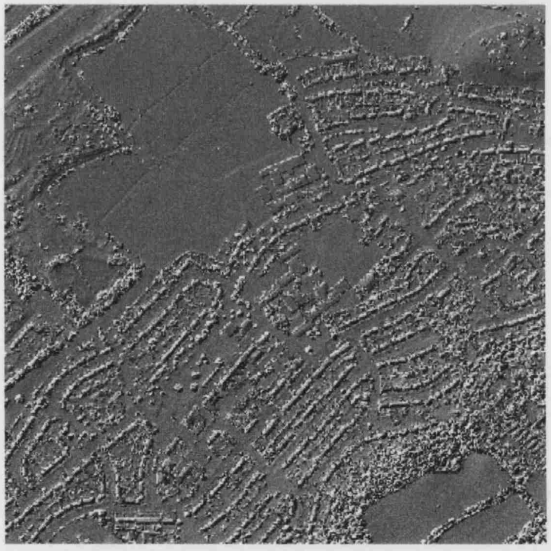

Figure D.5: Bicubic 1m DSM

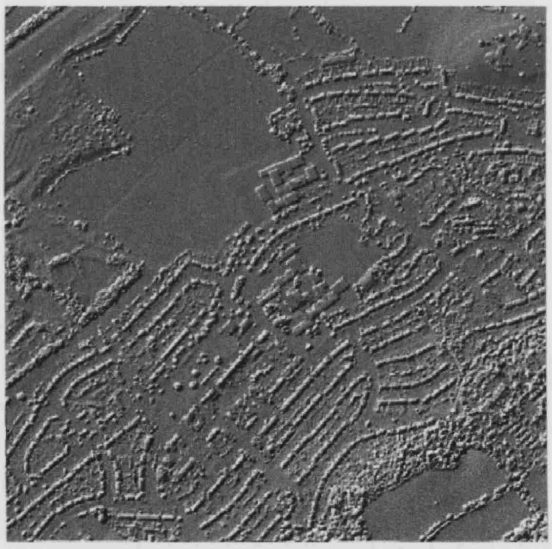

Figure D.6: Bicubic 2m DSM

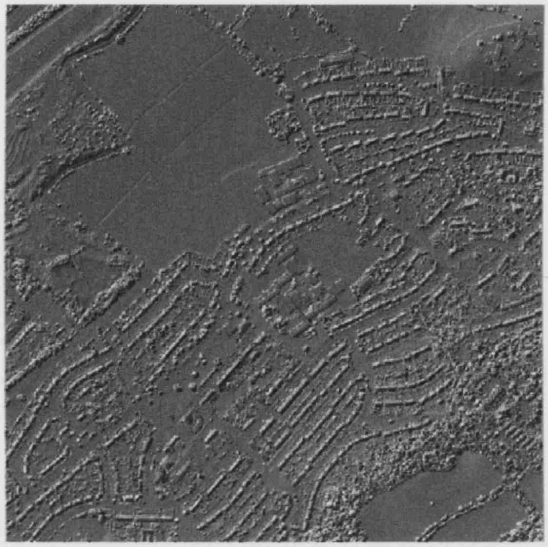

Figure D.7: Nearest Neighbour 0.5m DSM

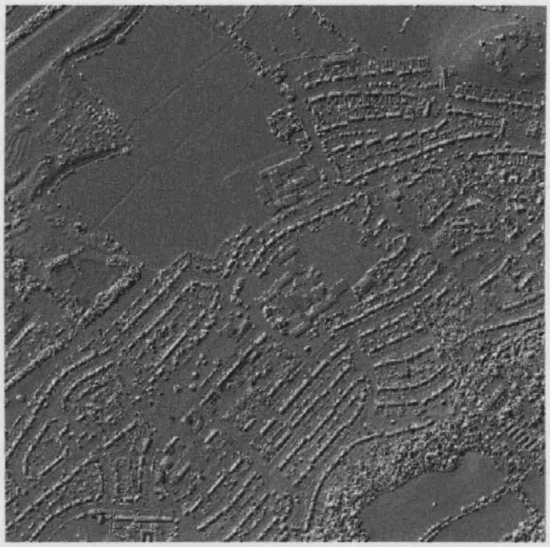

Figure D.8: Nearest Neighbour 1m DSM

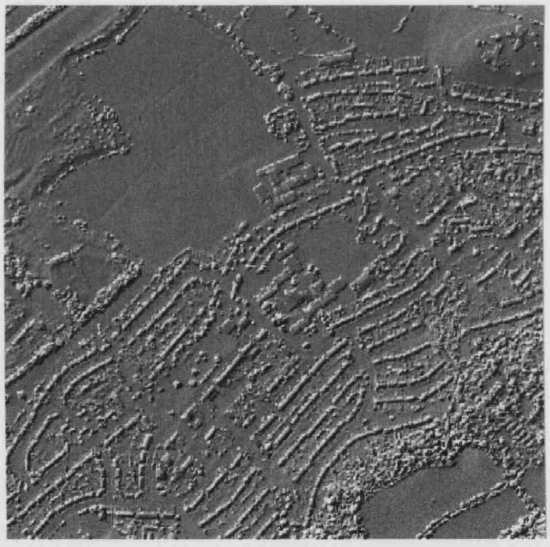

Figure D.9: Nearest Neighbour 2m DSM

Appendix E

Scripts for Geometric Method

E.1 Binary Classification

```
PAGESIZE 6, 8 INCHES;
CELLSIZE MINIMUM;
PRINTERPAGESIZE 8.5, 11;
MARGINS 0.5, 0.5, 0.5, 0.5;
ORIENTATION PORTRAIT;
PRINTSCALE 100;
WINDOW UNION;
PROJECTION DEFAULT;
AOI NONE;
OPTIMIZE NO;
RASTER {
ID 1;
TITLE "n1_classification";
POSITION 2.75556, 1.04444;
FILENAME "u:/users/classification.img";
INTERPOLATION NEAREST;
THEMATIC;
DATATYPE UNSIGNED8;
DECLARE "Integer";
COMPRESSION UNCOMPRESSED;
COORDINATES MAP;
AOI NONE;
RECODE NO;
CHILD 3;
}
RASTER {
ID 2;
```

```
TITLE "n2_binaryclass_linear2m";
POSITION 2.8, 4.67778;
FILENAME "u:/users/binaryclass_linear2m.img";
DELETEFILE;
INTERPOLATION NEAREST;
ATHEMATIC;
DATATYPE UNSIGNED8;
DECLARE "Integer";
COMPRESSION UNCOMPRESSED;
COORDINATES MAP;
RECODE NO;
}
FUNCTION {
ID 3;
TITLE "EITHER 1 IF ";
POSITION 2.77778, 2.75556;
VALUE "EITHER 1 IF ( n1_classification  <  2 ) OR 0 OTHERWISE ";
AREA UNION;
CHILD 2;
}
```

E.2 Example script for cookie-cutting the image

```
PAGESIZE 6, 8 INCHES;
CELLSIZE MINIMUM;
PRINTERPAGESIZE 8.5, 11;
MARGINS 0.5, 0.5, 0.5, 0.5;
ORIENTATION PORTRAIT;
PRINTSCALE 100;
WINDOW UNION;
PROJECTION DEFAULT;
AOI NONE;
OPTIMIZE NO;
RASTER {
 ID 1;
 TITLE "n1_binaryclass_linear2m";
 POSITION 1.5, 0.733333;
 FILENAME "u:/users/binaryclass_linear2m.img";
 INTERPOLATION NEAREST;
 ATHEMATIC;
 DATATYPE UNSIGNED8;
 DECLARE "Integer";
 COMPRESSION UNCOMPRESSED;
```

```
 COORDINATES MAP;
 AOI NONE;
 RECODE NO;
 CHILD 3;
}
RASTER {
 ID 2;
 TITLE "n2_linear2m";
 POSITION 3.92222, 0.7;
 FILENAME "u:/users/linear2m.img";
 INTERPOLATION NEAREST;
 ATHEMATIC;
 DATATYPE FLOAT;
 DECLARE "Float";
 COMPRESSION UNCOMPRESSED;
 COORDINATES MAP;
 AOI NONE;
 RECODE NO;
 CHILD 3;
}
FUNCTION {
  ID 3;
TITLE "n1_binaryclass_linear2m";
POSITION 2.72222, 2.37778;
  VALUE "n1_binaryclass_linear2m * n2_linear2m";
AREA UNION;
CHILD 4;
}
RASTER {
ID 4;
TITLE "n4_gaps_linear2m";
POSITION 2.74444, 4.03333;
FILENAME "u:/users/gaps_linear2m.img";
DELETEFILE;
INTERPOLATION NEAREST;
ATHEMATIC;
DATATYPE DOUBLE;
DECLARE "Float";
COMPRESSION UNCOMPRESSED;
COORDINATES MAP;
RECODE NO;
}
```

E.3 Script for removing the object halos

```
PAGESIZE 6, 8 INCHES;
CELLSIZE MINIMUM;
PRINTERPAGESIZE 8.5, 11;
MARGINS 0.5, 0.5, 0.5, 0.5;
ORIENTATION PORTRAIT;
PRINTSCALE 100;
WINDOW UNION;
PROJECTION DEFAULT;
AOI NONE;
OPTIMIZE NO;
RASTER {
ID 1;
TITLE "n1_gaps_linear2m";
POSITION 1.3, 0.877778;
FILENAME "u:/users/gaps_linear2m.img";
INTERPOLATION NEAREST;
ATHEMATIC;
DATATYPE DOUBLE;
DECLARE "Float";
COMPRESSION UNCOMPRESSED;
COORDINATES MAP;
AOI NONE;
RECODE NO;
CHILD 3;
}
RASTER {
ID 2;
TITLE "n2_halo_removal";
POSITION 2.57778, 4;
FILENAME "u:/users/halo_removal.img";
DELETEFILE;
INTERPOLATION NEAREST;
ATHEMATIC;
DATATYPE DOUBLE;
DECLARE "Float";
COMPRESSION UNCOMPRESSED;
COORDINATES MAP;
RECODE NO;
}
FUNCTION {
ID 3;
TITLE "FOCAL MIN ";
POSITION 2.56667, 2.36667;
```

```
VALUE "FOCAL MIN (  n1_gaps_linear2m ,  n4_Custom_Float ) ";
AREA UNION;
CHILD 2;
}
MATRIX {
ID 4;
TITLE "n4_Custom_Float";
POSITION 3.72222, 0.844444;
SIZE 7, 7;
DATATYPE DOUBLE;
VALUE   1, 1, 1, 0, 0, 0, 0,
1, 1, 1, 0, 0, 0, 0,
1, 1, 1, 0, 0, 0, 0,
0, 0, 0, 0, 0, 0, 0,
0, 0, 0, 0, 0, 0, 0,
0, 0, 0, 0, 0, 0, 0,
0, 0, 0, 0, 0, 0, 0;
NORMALIZE YES;
CHILD 3;
}
```

E.4 Script for interpolating across the gaps to create the continuous surface

```
PAGESIZE 6.17778, 15.1333 INCHES;
CELLSIZE MINIMUM;
PRINTERPAGESIZE 8.5, 11;
MARGINS 0.5, 0.5, 0.5, 0.5;
ORIENTATION PORTRAIT;
PRINTSCALE 100;
WINDOW UNION;
PROJECTION DEFAULT;
AOI NONE;
OPTIMIZE NO;
RASTER {
ID 1;
TITLE "n1_halo_removal";
POSITION 1.83334, 0.366666;
FILENAME "u:/users/halo_removal.img";
INTERPOLATION NEAREST;
ATHEMATIC;
DATATYPE DOUBLE;
DECLARE "Float";
```

```
COMPRESSION UNCOMPRESSED;
COORDINATES MAP;
AOI NONE;
RECODE NO;
CHILD 15;
}
FUNCTION {
ID 6;
TITLE "FOCAL MEAN ";
POSITION 2.85555, 3.71111;
VALUE "FOCAL MEAN ( n14_memory , n13_Custom_Float ,
IGNORE_VALUE 0 , APPLY_AT_VALUE 0 ) ";
AREA UNION;
CHILD 10;
}
RASTER {
ID 10;
TITLE "n10_memory";
POSITION 2.87778, 5.13333;
TEMPFILE;
NEWFILE;
INTERPOLATION NEAREST;
ATHEMATIC;
DATATYPE FLOAT;
DECLARE "Integer";
COMPRESSION UNCOMPRESSED;
COORDINATES MAP;
RECODE NO;
CHILD 17;
}
MATRIX {
ID 11;
TITLE "n11_Custom_Float";
POSITION 4.97778, 13.9444;
SIZE 21, 21;
DATATYPE DOUBLE;
VALUE   1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1,
1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1,
1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1,
1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1,
1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1,
1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1,
1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1,
1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1,
1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1,
1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1,
```

```
1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1,
1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1,
1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1,
1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1,
1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1,
1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1,
1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1,
1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1,
1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1,
1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1;
NORMALIZE YES;
}
MATRIX {
ID 12;
TITLE "n12_Custom_Float";
POSITION 5.17778, 14.1333;
SIZE 21, 21;
DATATYPE DOUBLE;
VALUE   1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1,
1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1,
1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1,
1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1,
1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1,
1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1,
1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1,
1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1,
1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1,
1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1,
1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1,
1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1,
1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1,
1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1,
1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1,
1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1,
1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1,
1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1,
1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1,
1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1;
NORMALIZE YES;
}
MATRIX {
ID 13;
TITLE "n13_Custom_Float";
POSITION 4.03333, 2.34444;
```

```
SIZE 15, 15;
DATATYPE DOUBLE;
VALUE   1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1,
1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1,
1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1,
1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1,
1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1,
1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1,
1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1,
1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1,
1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1,
1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1,
1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1,
1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1,
1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1,
1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1;
NORMALIZE YES;
CHILD 6;
}
RASTER {
ID 14;
TITLE "n14_memory";
POSITION 1.9, 2.65556;
TEMPFILE;
NEWFILE;
INTERPOLATION NEAREST;
ATHEMATIC;
DATATYPE FLOAT;
DECLARE "Integer";
COMPRESSION UNCOMPRESSED;
COORDINATES MAP;
RECODE NO;
CHILD 6;
}
FUNCTION {
ID 15;
TITLE " n1_halo_removal";
  POSITION 1.83333, 1.54444;
VALUE "n1_halo_removal * 1000";
AREA UNION;
CHILD 14;
}
MATRIX {
ID 16;
TITLE "n16_Custom_Float";
```

```
POSITION 4.67778, 5.78889;
SIZE 10, 10;
DATATYPE DOUBLE;
VALUE   1, 1, 1, 1, 1, 1, 1, 1, 1, 1,
1, 1, 1, 1, 1, 1, 1, 1, 1, 1,
1, 1, 1, 1, 1, 1, 1, 1, 1, 1,
1, 1, 1, 1, 1, 1, 1, 1, 1, 1,
1, 1, 1, 1, 1, 1, 1, 1, 1, 1,
1, 1, 1, 1, 1, 1, 1, 1, 1, 1,
1, 1, 1, 1, 1, 1, 1, 1, 1, 1,
1, 1, 1, 1, 1, 1, 1, 1, 1, 1,
1, 1, 1, 1, 1, 1, 1, 1, 1, 1,
1, 1, 1, 1, 1, 1, 1, 1, 1, 1;
NORMALIZE YES;
CHILD 17;
}
FUNCTION {
ID 17;
TITLE "FOCAL MEAN ";
POSITION 2.94444, 6.5;
VALUE "FOCAL MEAN ( n10_memory ,  n16_Custom_Float ,
IGNORE_VALUE 0 , APPLY_AT_VALUE 0 ) ";
AREA UNION;
CHILD 21;
}
RASTER {
ID 19;
TITLE "n19_final_dtm_linear2m";
POSITION 2.82222, 10.1889;
FILENAME "u:/users/final_dtm_linear2m.img";
NEWFILE;
INTERPOLATION NEAREST;
ATHEMATIC;
DATATYPE DOUBLE;
DECLARE "Float";
COMPRESSION UNCOMPRESSED;
COORDINATES MAP;
RECODE NO;
}
FUNCTION {
ID 20;
TITLE "n21_dtm_linear2m";
POSITION 4.11111, 8.9;
VALUE "n21_dtm_linear2m / 1000";
AREA UNION;
CHILD 19;
```

```
}
RASTER {
ID 21;
TITLE "n21_dtm_linear2m";
POSITION 2.95556, 7.84444;
FILENAME "u:/users/dtm_linear2m.img";
DELETEFILE;
INTERPOLATION NEAREST;
ATHEMATIC;
DATATYPE DOUBLE;
DECLARE "Float";
COMPRESSION UNCOMPRESSED;
COORDINATES MAP;
RECODE NO;
CHILD 20;
}
```

Appendix F

Example Parameter File for IRI Method

```
1-------------------------------------------------------------------------
 ThinOut  VERS 5.2.1 step nb.0                          04Oct 04 12:16:07
-------------------------------------------------------------------------
Input: C:\INPHO_~1\SCOP__~1\NRST1M~1\NEARES~1\nearest1m_.all
Output: C:\Inpho_Data\SCOP++_Projects\nrst1m_scop\nearest1m\step0.tho
Cell size:           2.000
Method:              Mean
Area:
  Left lower corner: 354000.000 178000.000
  Size:              1000.000 1000.000
Nb of input points:  1002001
Nb of output points: 251001

 END SCOP.ThinOut

1-------------------------------------------------------------------------
  Filter step nb.1                          04Oct 04 12:19:16
-------------------------------------------------------------------------
Input: C:\Inpho_Data\SCOP++_Projects\nrst1m_scop\nearest1m\step0.tho
DTM: C:\Inpho_Data\SCOP++_Projects\nrst1m_scop\nearest1m\step1.dtm
Output ground: C:\Inpho_Data\SCOP++_Projects\nrst1m_scop\nearest1m\
step1.grd
Output vegetation: C:\Inpho_Data\SCOP++_Projects\nrst1m_scop\nearest1m\
step1.veg


   DIGITAL ELEVATION MODEL
```

```
      MAP SHEET


   POSITION AND SIZE
        LEFT LOWER CORNER          EAST  .........      354000.00
                                   NORTH .........      178000.00
        EXTENSION                  EAST  .........        1000.00
                                   NORTH .........        1000.00

   CHARACTERISTICS OF THE DEM
        GRID WIDTH                 EAST  .........           1.00
                                   NORTH .........           1.00
        NUMBER OF GRID LINES       EAST  .........             21
                                   NORTH .........             21
        NUMBER OF INTERPOLATED COMPUTING UNITS ....           2500
        NUMBER OF STORED GRID POINTS      .........        1102500
        NUMBER OF GRID INTERSECTIONS      .........              0

  INFORMATION ABOUT THE INTERPOLATION
        LINEAR PREDICTION
        NUMBER OF REFERENCE POINTS GIVEN .........         251001
              SINGLE POINTS              ..........         251001
              HIGHS AND LOWS             ..........              0
              LINE POINTS                ..........              0
        AVERAGE FILTER VALUES
              SINGLE POINTS              ..........           .831
              HIGHS AND LOWS             ..........           .000
              LINE POINTS                ..........           .000
        MAXIMUM FILTER VALUES
              SINGLE POINTS              ..........         17.479
              HIGHS AND LOWS             ..........           .000
              LINE POINTS                ..........           .000


1-------------------------------------------------------------------------
 ThinOut  VERS 5.2.1 step nb.2                          04Oct 04 12:19:30
-------------------------------------------------------------------------
Input: C:\Inpho_Data\SCOP++_Projects\nrst1m_scop\nearest1m\step1.grd
Output: C:\Inpho_Data\SCOP++_Projects\nrst1m_scop\nearest1m\step2.tho
Cell size:           5.000
Method:              Lowest
Area:
  Left lower corner: 354000.000 178000.000
  Size:              1000.000 1000.000
Nb of input points:  231024
Nb of output points: 40232
```

```
 END SCOP.ThinOut

1-------------------------------------------------------------------------
  Filter step nb.3                          04Oct 04 12:20:13
-------------------------------------------------------------------------
Input: C:\Inpho_Data\SCOP++_Projects\nrst1m_scop\nearest1m\step2.tho
DTM: C:\Inpho_Data\SCOP++_Projects\nrst1m_scop\nearest1m\step3.dtm
Output ground: C:\Inpho_Data\SCOP++_Projects\nrst1m_scop\nearest1m\
step3.grd
Output vegetation: C:\Inpho_Data\SCOP++_Projects\nrst1m_scop\nearest1m\
step3.veg


   DIGITAL ELEVATION MODEL
        MAP SHEET


   POSITION AND SIZE
        LEFT LOWER CORNER          EAST  .........      354000.00
                                   NORTH .........      178000.00
        EXTENSION                  EAST  .........        1000.00
                                   NORTH .........        1000.00

   CHARACTERISTICS OF THE DEM
        GRID WIDTH                 EAST  .........           5.00
                                   NORTH .........           5.00
        NUMBER OF GRID LINES       EAST  .........             16
                                   NORTH .........             16
        NUMBER OF INTERPOLATED COMPUTING UNITS ....            196
        NUMBER OF STORED GRID POINTS      .........          50176
        NUMBER OF GRID INTERSECTIONS      .........              0

  INFORMATION ABOUT THE INTERPOLATION
        LINEAR PREDICTION
        NUMBER OF REFERENCE POINTS GIVEN .........          40232
               SINGLE POINTS            ..........          40232
               HIGHS AND LOWS           ..........              0
               LINE POINTS              ..........              0
        AVERAGE FILTER VALUES
               SINGLE POINTS            ..........          1.666
               HIGHS AND LOWS           ..........           .000
               LINE POINTS              ..........           .000
        MAXIMUM FILTER VALUES
               SINGLE POINTS            ..........         24.422
               HIGHS AND LOWS           ..........           .000
```

```
            LINE POINTS                ..........          .000


1-------------------------------------------------------------------------
  Filter step nb.4                          04Oct 04 12:23:27
-------------------------------------------------------------------------
Input: C:\Inpho_Data\SCOP++_Projects\nrst1m_scop\nearest1m\step3.grd
DTM: C:\Inpho_Data\SCOP++_Projects\nrst1m_scop\nearest1m\step4.dtm
Output ground: C:\Inpho_Data\SCOP++_Projects\nrst1m_scop\nearest1m\
step4.grd
Output vegetation: C:\Inpho_Data\SCOP++_Projects\nrst1m_scop\nearest1m\
step4.veg


   DIGITAL ELEVATION MODEL
        MAP SHEET


   POSITION AND SIZE
        LEFT LOWER CORNER          EAST  ..........      354000.00
                                   NORTH ..........      178000.00
        EXTENSION                  EAST  ..........        1000.00
                                   NORTH ..........        1000.00

   CHARACTERISTICS OF THE DEM
        GRID WIDTH                 EAST  .........          5.00
                                   NORTH .........          5.00
        NUMBER OF GRID LINES       EAST  .........            21
                                   NORTH .........            21
        NUMBER OF INTERPOLATED COMPUTING UNITS ....           100
        NUMBER OF STORED GRID POINTS      .........         44100
        NUMBER OF GRID INTERSECTIONS      .........             0

  INFORMATION ABOUT THE INTERPOLATION
        LINEAR PREDICTION
        NUMBER OF REFERENCE POINTS GIVEN ..........         34284
               SINGLE POINTS             ..........         34284
               HIGHS AND LOWS            ..........             0
               LINE POINTS               ..........             0
        AVERAGE FILTER VALUES
               SINGLE POINTS             ..........          .208
               HIGHS AND LOWS            ..........          .000
               LINE POINTS               ..........          .000
        MAXIMUM FILTER VALUES
               SINGLE POINTS             ..........         3.559
               HIGHS AND LOWS            ..........          .000
```

```
            LINE POINTS                .....          .000


1-------------------------------------------------------------------------
  Interpolation step nb.5                              04Oct 04 12:23:41
-------------------------------------------------------------------------
Input: C:\Inpho_Data\SCOP++_Projects\nrst1m_scop\nearest1m\step4.grd
DTM: C:\Inpho_Data\SCOP++_Projects\nrst1m_scop\nearest1m\step5.dtm


   DIGITAL ELEVATION MODEL
        MAP SHEET


   POSITION AND SIZE
        LEFT LOWER CORNER          EAST  .....      354000.00
                                   NORTH .....      178000.00
        EXTENSION                  EAST  .....        1000.00
                                   NORTH .....        1000.00

   CHARACTERISTICS OF THE DEM
        GRID WIDTH                 EAST  .....           2.50
                                   NORTH .....           2.50
        NUMBER OF GRID LINES       EAST  .....             21
                                   NORTH .....             21
        NUMBER OF INTERPOLATED COMPUTING UNITS ....        400
        NUMBER OF STORED GRID POINTS      .....       176400
        NUMBER OF GRID INTERSECTIONS      .....            0

  INFORMATION ABOUT THE INTERPOLATION
        LINEAR PREDICTION
        NUMBER OF REFERENCE POINTS GIVEN .....         31042
              SINGLE POINTS          .....           31042
              HIGHS AND LOWS         .....               0
              LINE POINTS            .....               0
        AVERAGE FILTER VALUES
              SINGLE POINTS          .....            .095
              HIGHS AND LOWS         .....            .000
              LINE POINTS            .....            .000
        MAXIMUM FILTER VALUES
              SINGLE POINTS          .....           1.081
              HIGHS AND LOWS         .....            .000
              LINE POINTS            .....            .000


1-------------------------------------------------------------------------
```

```
 SortOut  VERS 5.2.1 step nb.6                          04Oct 04 12:25:09
-------------------------------------------------------------------------
Input: C:\INPHO_~1\SCOP__~1\NRST1M~1\NEARES~1\nearest1m_.all
DTM: C:\Inpho_Data\SCOP++_Projects\nrst1m_scop\nearest1m\step5.dtm
Output: C:\Inpho_Data\SCOP++_Projects\nrst1m_scop\nearest1m\step6.sog
Lower distance:      -3.000
Upper distance:      3.000
Nb of input points:  1002001
Nb of output points: 808540

 END SCOP.SortOut

1-------------------------------------------------------------------------
 ThinOut  VERS 5.2.1 step nb.7                          04Oct 04 12:25:56
-------------------------------------------------------------------------
Input: C:\Inpho_Data\SCOP++_Projects\nrst1m_scop\nearest1m\step6.sog
Output: C:\Inpho_Data\SCOP++_Projects\nrst1m_scop\nearest1m\step7.tho
Cell size:           2.000
Method:              Lowest
Area:
  Left lower corner: 354000.000 178000.000
  Size:              1000.000 1000.000
Nb of input points:  808540
Nb of output points: 217346

 END SCOP.ThinOut

1-------------------------------------------------------------------------
  Filter step nb.8                          04Oct 04 12:41:04
-------------------------------------------------------------------------
Input: C:\Inpho_Data\SCOP++_Projects\nrst1m_scop\nearest1m\step7.tho
DTM: C:\Inpho_Data\SCOP++_Projects\nrst1m_scop\nearest1m\step8.dtm
Output ground: C:\Inpho_Data\SCOP++_Projects\nrst1m_scop\nearest1m\
step8.grd
Output vegetation: C:\Inpho_Data\SCOP++_Projects\nrst1m_scop\nearest1m\
step8.veg


   DIGITAL ELEVATION MODEL
        MAP SHEET


   POSITION AND SIZE
        LEFT LOWER CORNER          EAST  .........      354000.00
                                   NORTH .........      178000.00
        EXTENSION                  EAST  .........        1000.00
```

```
                                NORTH .....          1000.00

   CHARACTERISTICS OF THE DEM
        GRID WIDTH              EAST   .....            2.00
                                NORTH  .....            2.00
        NUMBER OF GRID LINES    EAST   .....              21
                                NORTH  .....              21
        NUMBER OF INTERPOLATED COMPUTING UNITS ....      625
        NUMBER OF STORED GRID POINTS      .....       275625
        NUMBER OF GRID INTERSECTIONS      .....            0

  INFORMATION ABOUT THE INTERPOLATION
        LINEAR PREDICTION
        NUMBER OF REFERENCE POINTS GIVEN .....       217346
               SINGLE POINTS             .....       217346
               HIGHS AND LOWS            .....            0
               LINE POINTS               .....            0
        AVERAGE FILTER VALUES
               SINGLE POINTS             .....         .320
               HIGHS AND LOWS            .....         .000
               LINE POINTS               .....         .000
        MAXIMUM FILTER VALUES
               SINGLE POINTS             .....        8.437
               HIGHS AND LOWS            .....         .000
               LINE POINTS               .....         .000


1-------------------------------------------------------------------------
  Interpolation step nb.9                          04Oct 04 12:41:32
-------------------------------------------------------------------------
Input: C:\Inpho_Data\SCOP++_Projects\nrst1m_scop\nearest1m\step8.grd
DTM: C:\Inpho_Data\SCOP++_Projects\nrst1m_scop\nearest1m\step9.dtm


   DIGITAL ELEVATION MODEL
        MAP SHEET


   POSITION AND SIZE
        LEFT LOWER CORNER       EAST   .....     354000.00
                                NORTH  .....     178000.00
        EXTENSION               EAST   .....       1000.00
                                NORTH  .....       1000.00

   CHARACTERISTICS OF THE DEM
        GRID WIDTH              EAST   .....          1.00
```

```
                                NORTH  .........          1.00
        NUMBER OF GRID LINES    EAST   .........           11
                                NORTH  .........           11
        NUMBER OF INTERPOLATED COMPUTING UNITS ....       9996
        NUMBER OF STORED GRID POINTS      .........    1209516
        NUMBER OF GRID INTERSECTIONS      .........          0

  INFORMATION ABOUT THE INTERPOLATION
        LINEAR PREDICTION
        NUMBER OF REFERENCE POINTS GIVEN ..........     199314
               SINGLE POINTS             ..........     199314
               HIGHS AND LOWS            ..........          0
               LINE POINTS               ..........          0
        AVERAGE FILTER VALUES
               SINGLE POINTS             ..........       .089
               HIGHS AND LOWS            ..........       .000
               LINE POINTS               ..........       .000
        MAXIMUM FILTER VALUES
               SINGLE POINTS             ..........      1.108
               HIGHS AND LOWS            ..........       .000
               LINE POINTS               ..........       .000


1-------------------------------------------------------------------------
 SortOut  VERS 5.2.1 step nb.10                          04Oct 04 12:43:00
-------------------------------------------------------------------------
Input: C:\INPHO_~1\SCOP__~1\NRST1M~1\NEARES~1\nearest1m_.all
DTM: C:\Inpho_Data\SCOP++_Projects\nrst1m_scop\nearest1m\step9.dtm
Output:              C:\Inpho_Data\SCOP++_Projects\nrst1m_scop\nearest1m\
step10.sog
Lower distance:      -3.000
Upper distance:      1.500
Nb of input points:  1002001
Nb of output points: 747550

 END SCOP.SortOut

1-------------------------------------------------------------------------
  Filter step nb.11                          04Oct 04 12:46:15
-------------------------------------------------------------------------
Input: C:\Inpho_Data\SCOP++_Projects\nrst1m_scop\nearest1m\step10.sog
DTM: C:\Inpho_Data\SCOP++_Projects\nrst1m_scop\nearest1m\step11.dtm
Output ground: C:\Inpho_Data\SCOP++_Projects\nrst1m_scop\nearest1m\
step11.grd
Output vegetation: C:\Inpho_Data\SCOP++_Projects\nrst1m_scop\nearest1m\
step11.veg
```

```
DIGITAL ELEVATION MODEL
     MAP SHEET


POSITION AND SIZE
     LEFT LOWER CORNER          EAST  .........      354000.00
                                NORTH .........      178000.00
     EXTENSION                  EAST  .........        1000.00
                                NORTH .........        1000.00

CHARACTERISTICS OF THE DEM
     GRID WIDTH                 EAST  .........           1.00
                                NORTH .........           1.00
     NUMBER OF GRID LINES       EAST  .........              9
                                NORTH .........              9
     NUMBER OF INTERPOLATED COMPUTING UNITS ....          15614
     NUMBER OF STORED GRID POINTS      .........        1264734
     NUMBER OF GRID INTERSECTIONS      .........              0

INFORMATION ABOUT THE INTERPOLATION
     LINEAR PREDICTION
     NUMBER OF REFERENCE POINTS GIVEN .........         747550
           SINGLE POINTS              ..........         747550
           HIGHS AND LOWS             ..........              0
           LINE POINTS                ..........              0
     AVERAGE FILTER VALUES
           SINGLE POINTS              ..........           .210
           HIGHS AND LOWS             ..........           .000
           LINE POINTS                ..........           .000
     MAXIMUM FILTER VALUES
           SINGLE POINTS              ..........          4.958
           HIGHS AND LOWS             ..........           .000
           LINE POINTS                ..........           .000


1-------------------------------------------------------------------------
  Filter step nb.12                          04Oct 04 12:51:51
-------------------------------------------------------------------------
Input: C:\Inpho_Data\SCOP++_Projects\nrst1m_scop\nearest1m\step11.grd
DTM: C:\Inpho_Data\SCOP++_Projects\nrst1m_scop\nearest1m\step12.dtm
Output ground: C:\Inpho_Data\SCOP++_Projects\nrst1m_scop\nearest1m\
step12.grd
Output vegetation: C:\Inpho_Data\SCOP++_Projects\nrst1m_scop\nearest1m\
step12.veg
```

```
   DIGITAL ELEVATION MODEL
        MAP SHEET


   POSITION AND SIZE
        LEFT LOWER CORNER          EAST  .........      354000.00
                                   NORTH .........      178000.00
        EXTENSION                  EAST  .........        1000.00
                                   NORTH .........        1000.00

   CHARACTERISTICS OF THE DEM
        GRID WIDTH                 EAST  .........           1.00
                                   NORTH .........           1.00
        NUMBER OF GRID LINES       EAST  .........             10
                                   NORTH .........             10
        NUMBER OF INTERPOLATED COMPUTING UNITS ....          12535
        NUMBER OF STORED GRID POINTS      .........        1253500
        NUMBER OF GRID INTERSECTIONS      .........              0

  INFORMATION ABOUT THE INTERPOLATION
        LINEAR PREDICTION
        NUMBER OF REFERENCE POINTS GIVEN ..........         728495
              SINGLE POINTS           ..........         728495
              HIGHS AND LOWS          ..........              0
              LINE POINTS             ..........              0
        AVERAGE FILTER VALUES
              SINGLE POINTS           ..........           .115
              HIGHS AND LOWS          ..........           .000
              LINE POINTS             ..........           .000
        MAXIMUM FILTER VALUES
              SINGLE POINTS           ..........          2.435
              HIGHS AND LOWS          ..........           .000
              LINE POINTS             ..........           .000


1-------------------------------------------------------------------------
  Interpolation step nb.13                          04Oct 04 12:54:07
-------------------------------------------------------------------------
Input: C:\Inpho_Data\SCOP++_Projects\nrst1m_scop\nearest1m\step12.grd
DTM: C:\Inpho_Data\SCOP++_Projects\nrst1m_scop\nearest1m\step13.dtm


   DIGITAL ELEVATION MODEL
        MAP SHEET
```

```
POSITION AND SIZE
     LEFT LOWER CORNER          EAST  .........      354000.00
                                NORTH .........      178000.00
     EXTENSION                  EAST  .........        1000.00
                                NORTH .........        1000.00

CHARACTERISTICS OF THE DEM
     GRID WIDTH                 EAST  .........           .50
                                NORTH .........           .50
     NUMBER OF GRID LINES       EAST  .........          21
                                NORTH .........          21
     NUMBER OF INTERPOLATED COMPUTING UNITS ....        9996
     NUMBER OF STORED GRID POINTS      .........     4408236
     NUMBER OF GRID INTERSECTIONS      .........           0

INFORMATION ABOUT THE INTERPOLATION
     LINEAR PREDICTION
     NUMBER OF REFERENCE POINTS GIVEN .........      644128
           SINGLE POINTS              ..........      644128
           HIGHS AND LOWS             ..........           0
           LINE POINTS                ..........           0
     AVERAGE FILTER VALUES
           SINGLE POINTS              ..........        .062
           HIGHS AND LOWS             ..........        .000
           LINE POINTS                ..........        .000
     MAXIMUM FILTER VALUES
           SINGLE POINTS              ..........        .802
           HIGHS AND LOWS             ..........        .000
           LINE POINTS                ..........        .000

           LINE POINTS                ..........        .000
```

Appendix G

The resultant DTMs from filtering using IRI algorithm

G.1 IRI Default method

Figure G.1: Hill shaded DTM created using raw input data at original point spacing resolution

Figure G.2: Height coloured DTM created using raw input data at original point spacing

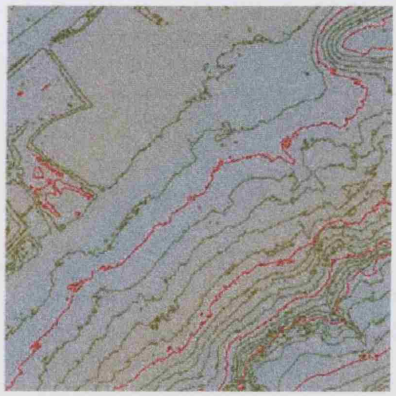

Figure G.3: Contour map of DTM created using raw input data at original point spacing resolution

Figure G.4: Hillshaded DTM created using input data at bilinear gridded data at 0.5m resolution

Figure G.5: Height coloured DTM created using input data at bilinear gridded data at 0.5m resolution

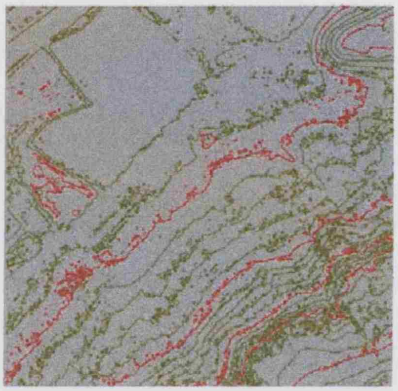

Figure G.6: Contour map of DTM created using input data at bilinear gridded data at 0.5m resolution

Figure G.7: Hillshaded DTM created using bilinear input gridded data at 1m resolution

Figure G.8: Height coloured DTM created using input data at bilinear gridded data at 1m resolution

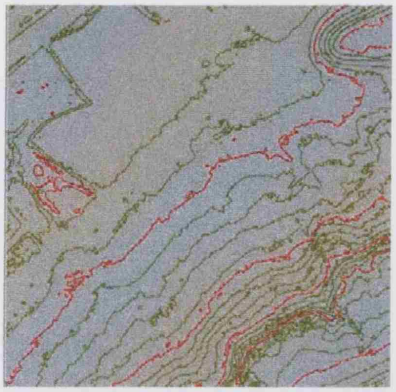

Figure G.9: Contour map of DTM created using bilinear input gridded data at 1m resolution

Figure G.10: Hillshaded DTM created using input data at bilinear gridded data at 2m resolution

Figure G.11: Height coloured DTM created using input data at bilinear gridded data at 2m resolution

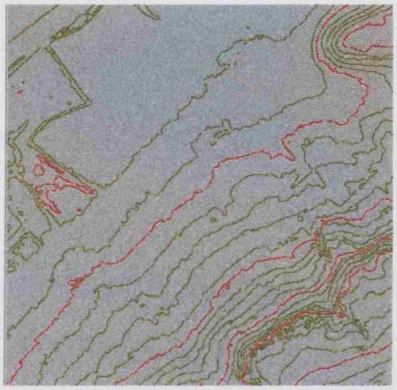

Figure G.12: Contour map of DTM created using input data at bilinear gridded data at 2m resolution

Figure G.13: Hillshaded DTM created using bicubic input gridded data at 0.5m resolution

Figure G.14: Height coloured DTM created using bicubic input gridded data at 0.5m resolution

Figure G.15: Contour map of DTM created using bicubic input gridded data at 0.5m resolution

Figure G.16: Hillshaded DTM created using bicubic input gridded data at 1m resolution

Figure G.17: Height coloured DTM created using bicubic input gridded data at 1m resolution

Figure G.18: Contour map of DTM created using bicubic input gridded data at 1m resolution

Figure G.19: Hillshaded DTM created using bicubic input gridded data at 2m resolution

Figure G.20: Height coloured DTM created using bicubic input gridded data at 2m resolution

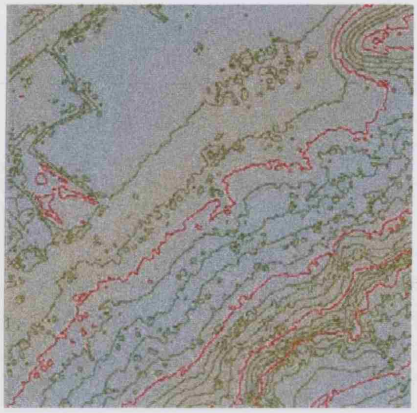

Figure G.21: Contour map of DTM created using bicubic input gridded data at 2m resolution

Figure G.22: Hillshaded DTM created using nearest neighbour input gridded data at 0.5m resolution

Figure G.23: Height coloured DTM created using nearest neighbour input gridded data at 0.5m resolution

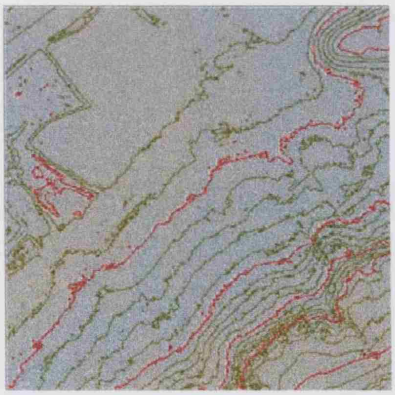

Figure G.24: Contour map of DTM created using nearest neighbour input gridded data at 0.5m resolution

Figure G.25: Hillshaded DTM created using nearest neighbour input gridded data at 1m resolution

Figure G.26: Height coloured DTM created using nearest neighbour input gridded data at 1m resolution

Figure G.27: Contour map of DTM created using nearest neighbour input gridded data at 1m resolution

Figure G.28: Hillshaded DTM created using nearest neighbour input gridded data at 2m resolution

Figure G.29: Height coloured DTM created using nearest neighbour input gridded data at 2m resolution

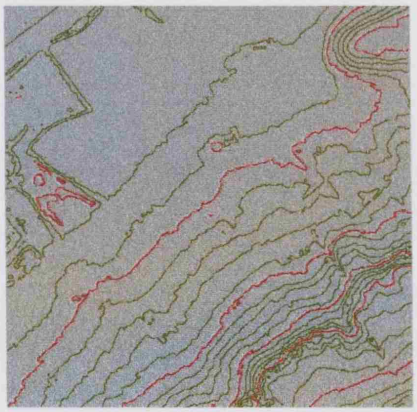

Figure G.30: Contour map of DTM created using nearest neighbour input gridded data at 2m resolution

G.2 IRI Adapted method

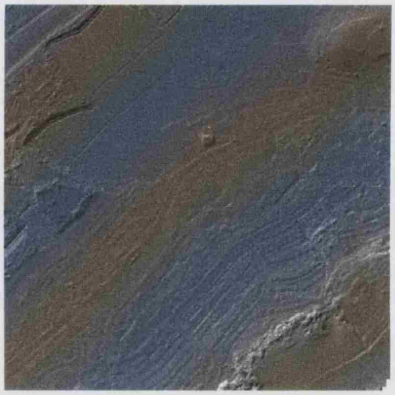

Figure G.31: Hill shaded DTM created using raw input data at original point spacing resolution

Figure G.32: Height coloured DTM created using raw input data at original point spacing resolution

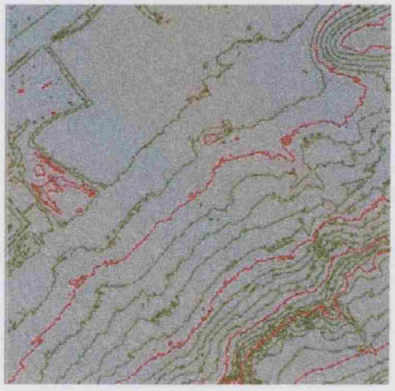

Figure G.33: Contour map of DTM created using raw input data at original point spacing resolution

Figure G.34: Hillshaded DTM created using input data at bilinear gridded data at 0.5m resolution

Figure G.35: Height coloured DTM created using input data at bilinear gridded data at 0.5m resolution

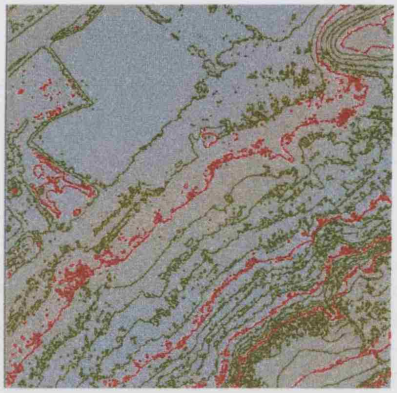

Figure G.36: Contour map of DTM created using input data at bilinear gridded data at 0.5m resolution

Figure G.37: Hillshaded DTM created using bilinear input gridded data at 1m resolution

Figure G.38: Height coloured DTM created using input data at bilinear gridded data at 1m resolution

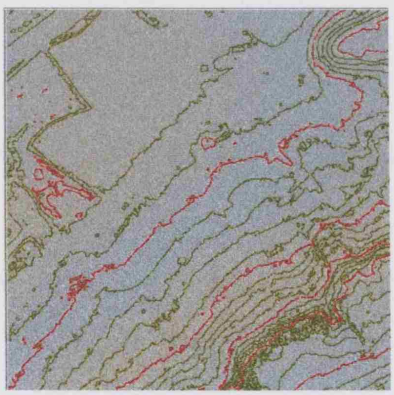

Figure G.39: Contour map of DTM created using bilinear input gridded data at 1m resolution

Figure G.40: Hillshaded DTM created using input data at bilinear gridded data at 2m resolution

Figure G.41: Height coloured DTM created using input data at bilinear gridded data at 2m resolution

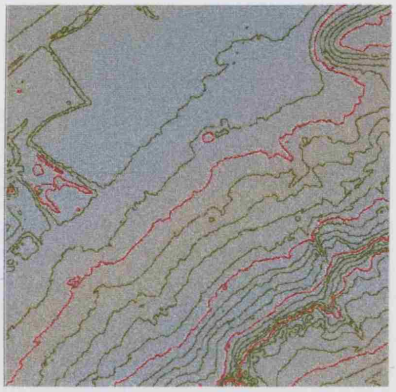

Figure G.42: Contour map of DTM created using input data at bilinear gridded data at 2m resolution

Figure G.43: Hillshaded DTM created using bicubic input gridded data at 0.5m resolution

Figure G.44: Height coloured DTM created using bicubic input gridded data at 0.5m resolution

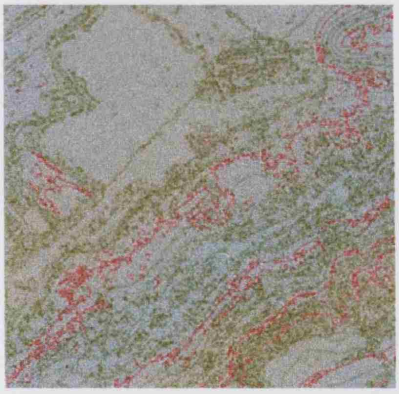

Figure G.45: Contour map of DTM created using bicubic input gridded data at 0.5m resolution

Figure G.46: Hillshaded DTM created using bicubic input gridded data at 1m resolution

Figure G.47: Height coloured DTM created using bicubic input gridded data at 1m resolution

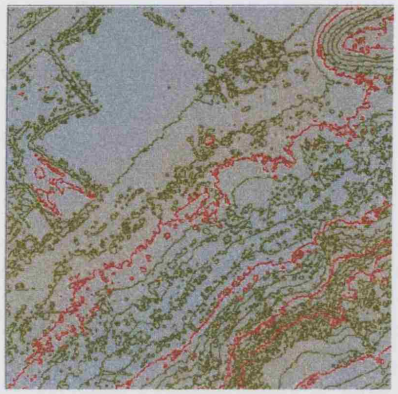

Figure G.48: Contour map of DTM created using bicubic input gridded data at 1m resolution

Figure G.49: Hillshaded DTM created using bicubic input gridded data at 2m resolution

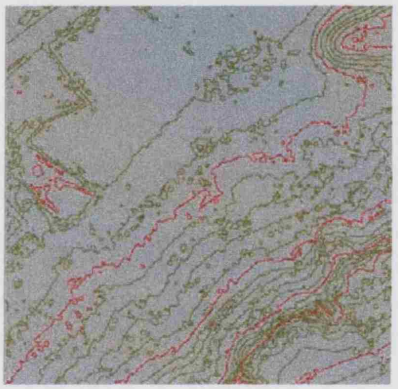

Figure G.50: Contour map of DTM created using bicubic input gridded data at 2m resolution

Figure G.51: Height coloured DTM created using bicubic input gridded data at 2m resolution

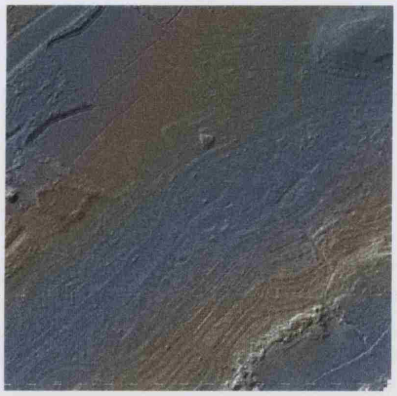

Figure G.52: Hillshaded DTM created using nearest neighbour input gridded data at 0.5m resolution

Figure G.53: Height coloured DTM created using nearest neighbour input gridded data at 0.5m resolution

Figure G.54: Contour map of DTM created using nearest neighbour input gridded data at 0.5m resolution

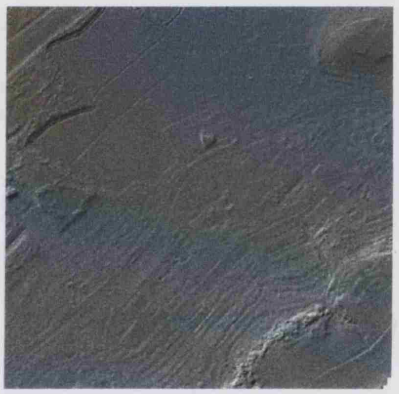

Figure G.55: Hillshaded DTM created using nearest neighbour input gridded data at 1m resolution

Figure G.56: Height coloured DTM created using nearest neighbour input gridded data at 1m resolution

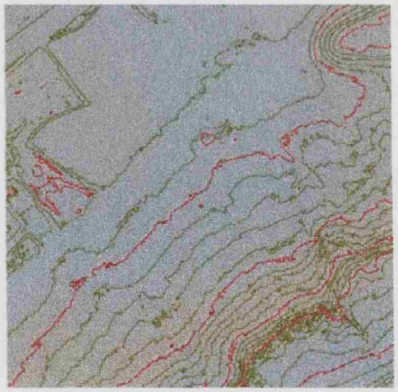

Figure G.57: Contour map of DTM created using nearest neighbour input gridded data at 1m resolution

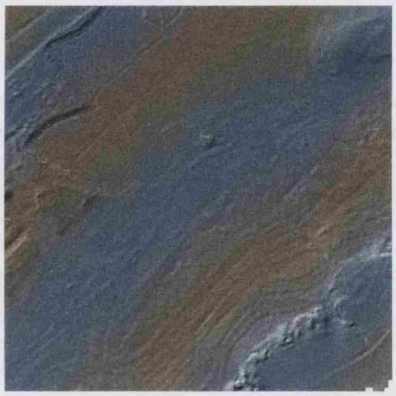

Figure G.58: Hillshaded DTM created using nearest neighbour input gridded data at 2m resolution

Figure G.59: Height coloured DTM created using nearest neighbour input gridded data at 2m resolution

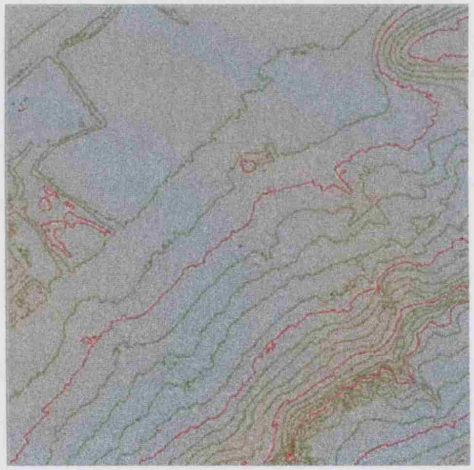

Figure G.60: Contour map of DTM created using nearest neighbour input gridded data at 2m resolution

Appendix H

Resultant DTMs and Associated Settings from ATM pilot study

H.1 Settings 1 - Terrain angle set at 5 degrees

H.1.1 Run 1

- Terrain angle: 5 degrees
- Iteration angle: 4 degrees
- Iteration distance: 1.2m

Note with the terrain angle set so low there is a problem in the more steeply sloping terrain in the ll corner

H.1.2 Run 2

- Terrain angle: 5 degrees
- Iteration angle: 3 degrees
- Iteration distance: 1.2m

Same problem in ll corner, also some clear visible differences in the filtering, these are obviously caused by change in the iteration angle

H.1.3 Run 3

- Terrain angle: 5 degrees
- Iteration angle: 4 degrees
- Iteration distance: 1.5m

H.1.4 Run 4

- Terrain angle: 5 degrees
- Iteration angle: 3 degrees
- Iteration distance: 1.5m

Figure H.1: Classification created using Settings 1 Run 1

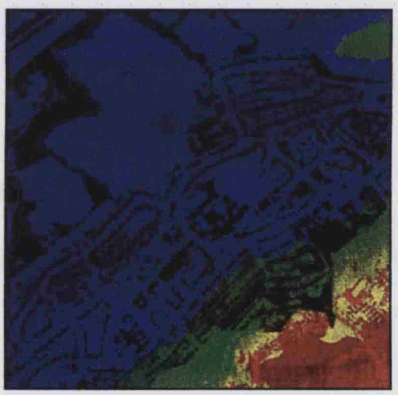

Figure H.2: Terrain points created using Settings 1 Run 1

Figure H.3: DTM created using Settings 1 Run 2

Figure H.4: Classification created using Settings 1 Run 2

Figure H.5: Terrain points created using Settings 1 Run 2

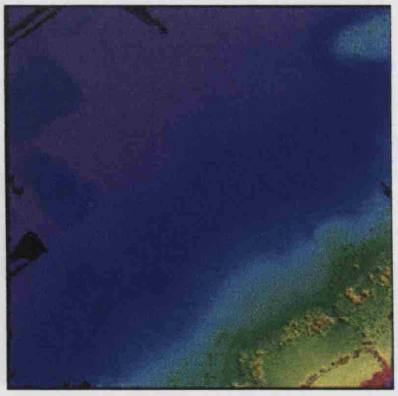

Figure H.6: DTM created using Settings 1 Run 2

Figure H.7: Classification created using Settings 1 Run 3

Figure H.8: Terrain points created using Settings 1 Run 3

Figure H.9: DTM created using Settings 1 Run 3

Figure H.10: Classification created using Settings 1 Run 3

Figure H.11: Terrain points created using Settings 1 Run 3

Figure H.12: DTM created using Settings 1 Run 3

H.2 Settings 2 - Terrain angle set at 10 degrees

H.2.1 Run 1

- Terrain angle: 10 degrees
- Iteration angle: 4 degrees
- Iteration distance: 1.2m

H.2.2 Run 2

- Terrain angle: 10 degrees
- Iteration angle: 3 degrees
- Iteration distance: 1.2m

H.2.3 Run 3

- Terrain angle: 10 degrees
- Iteration angle: 3 degrees
- Iteration distance: 1.5m

H.2.4 Run 4

- Terrain angle: 10 degrees
- Iteration angle: 4 degrees
- Iteration distance: 1.5m

Figure H.13: Classification created using Settings 2 Run 1

Figure H.14: Terrain points created using Settings 2 Run 1

Figure H.15: DTM created using Settings 2 Run 2

Figure H.16: Classification created using Settings 2 Run 2

Figure H.17: Terrain points created using Settings 2 Run 2

Figure H.18: DTM created using Settings 2 Run 2

Figure H.19: Classification created using Settings 2 Run 3

Figure H.20: Terrain points created using Settings 2 Run 3

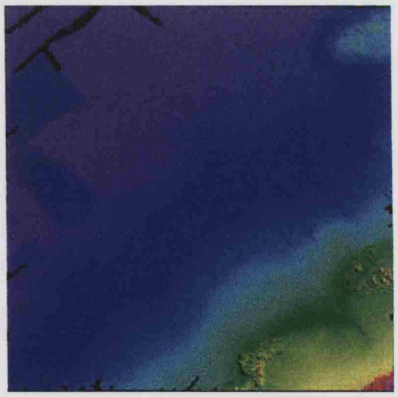

Figure H.21: DTM created using Settings 2 Run 3

Figure H.22: Classification created using Settings 2 Run 3

Figure H.23: Terrain points created using Settings 2 Run 3

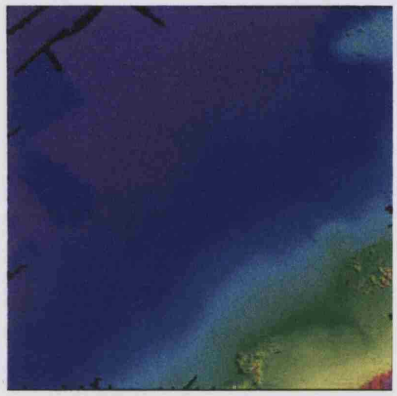

Figure H.24: DTM created using Settings 2 Run 3

H.3 Settings 3 - Terrain angle set at 15 degrees

H.3.1 Run 1

- Terrain angle: 15 degrees
- Iteration angle: 4 degrees
- Iteration distance: 1.2m

H.3.2 Run 2

- Terrain angle: 15 degrees
- Iteration angle: 3 degrees
- Iteration distance: 1.2m

H.3.3 Run 3

- Terrain angle: 15 degrees
- Iteration angle: 3 degrees
- Iteration distance: 1.5m

H.3.4 Run 4

- Terrain angle: 15 degrees
- Iteration angle: 4 degrees
- Iteration distance: 1.5m

Figure H.25: Classification created using Settings 3 Run 1

Figure H.26: Terrain points created using Settings 3 Run 1

Figure H.27: DTM created using Settings 3 Run 2

Figure H.28: Classification created using Settings 3 Run 2

Figure H.29: Terrain points created using Settings 3 Run 2

Figure H.30: DTM created using Settings 3 Run 2

Figure H.31: Classification created using Settings 3 Run 3

Figure H.32: Terrain points created using Settings 3 Run 3

Figure H.33: DTM created using Settings 3 Run 3

Figure H.34: Classification created using Settings 3 Run 3

Figure H.35: Terrain points created using Settings 3 Run 3

Figure H.36: DTM created using Settings 3 Run 3

H.4 Settings 4 - Terrain angle set at 20 degrees

H.4.1 Run 1

- Terrain angle: 20 degrees
- Iteration angle: 4 degrees
- Iteration distance: 1.2m

H.4.2 Run 2

- Terrain angle: 20 degrees
- Iteration angle: 3 degrees
- Iteration distance: 1.2m

H.4.3 Run 3

- Terrain angle: 20 degrees
- Iteration angle: 3 degrees
- Iteration distance: 1.5m

H.4.4 Run 4

- Terrain angle: 20 degrees
- Iteration angle: 4 degrees
- Iteration distance: 1.5m

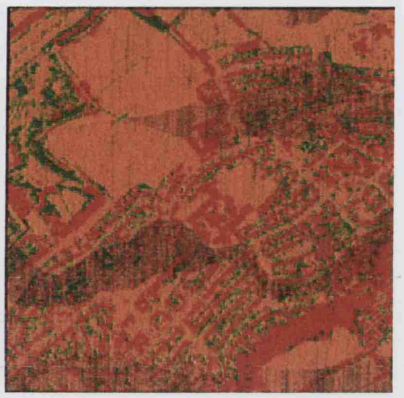

Figure H.37: Classification created using Settings 4 Run 1

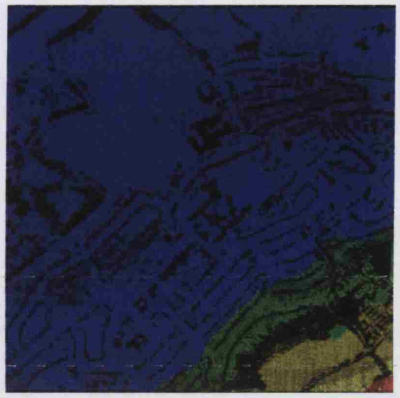

Figure H.38: Terrain points created using Settings 4 Run 1

Figure H.39: DTM created using Settings 4 Run 2

Figure H.40: Classification created using Settings 4 Run 2

Figure H.41: Terrain points created using Settings 4 Run 2

Figure H.42: DTM created using Settings 4 Run 2

Figure H.43: Classification created using Settings 4 Run 3

Figure H.44: Terrain points created using Settings 4 Run 3

Figure H.45: DTM created using Settings 4 Run 3

Figure H.46: Classification created using Settings 4 Run 3

Figure H.47: Terrain points created using Settings 4 Run 3

Figure H.48: DTM created using Settings 4 Run 3

Appendix I

Results from Repeat Experiments Using Four Research Assistants

Results of McNemar Test for Significant Difference

H_0 There is no signficant difference between original classification and the four research assistant classifications

H_1 There is a siginificant difference between original classification and research assistant classifications

Acccording to the McNemar Test z is defined as

$$z = \frac{f_{12} - f_{21}}{\sqrt{f_{12} + f_{21}}} \tag{I.1}$$

Where f_ij indicates the frequency of sites lying in confusion matrix element i, j (after Foody (2004))

I.1 Comparison with Research Assistant 1

For the comparison of the original classification and Research Assistant 1 classification the following contingency tables were compiled:

		Original classification		
		Ground	**Above ground**	
Research Assistant	**Ground**	68	28	96
David	**Above ground**	9	87	96
		77	115	192
				80.72916667

Using the McNemar equation, $calc - z$ yields a value of 3.12. This can be compared with the critical values for the McNemar test (as published in Siegel (1956) pg.249), using this table $tab - z$ is 3.84 at the 5% significance level with 1df. As $calc - z$ is less than $tab - z$ the null hypothesis is not rejected, indicating that there is no significant difference between the two classifications.

I.2 Comparison with Research Assistant 2

		Oringial classificaiton		
		Ground	**Above ground**	
Research Assistant	**Ground**	60	36	96
Doreen	**Above ground**	25	71	96
		63	129	192

$Calc - z = 1.41$

Since $Calcz$ is less than $Tabz$ do not reject H_0

		Original classification		
		Ground	Above ground	
Research Assistant	**Ground**	56	40	96
Will	**Above ground**	15	81	96
		71	121	192

I.3 Comparison with Research Assistant 3

$Calc - z = 3.37$

Since *Calcz* is less than *Tabz* do not reject H_0

I.4 Comparison with Research Assistant 4

		Original classification		
		Ground	Above ground	
Research Assistant	**Ground**	61	35	96
Jon	**Above Ground**	19	77	96
		80	112	192

$Calc - z = 2.18$

Since *Calcz* is less than *Tabz* do not reject H_0

Appendix J

Producer's Accuracy Results

Table J.1: Producer's Accuracy Results for Geometric Algorithm

DEFAULT		Producer's Accuracy %
Linear	2m	60.36
	1m	58.02
	0.5m	65.66
Nearest Neighbour	2m	62.16
	1m	76.92
	0.5m	71.43
Cubic	2m	63.71
	1m	66.38
	0.5m	60.71

ADAPTED		Producer's Accuracy %
Linear	2m	78.72
	1m	81.71
	0.5m	74.47
Nearest Neighbour	2m	96.05
	1m	93.59
	0.5m	97.37
Cubic	2m	78.26
	1m	80.65
	0.5m	73.11

Table J.2: Producer's Accuracy Results for ATM Algorithm

DEFAULT		Producer's Accuracy %
Linear	2m	63.25
	1m	72.03
	0.5m	68.75
Nearest Neighbour	2m	74.53
	1m	72.57
	0.5m	69.64
Cubic	2m	65.87
	1m	54.84
	0.5m	46.08
Raw data	1m	69.75

ADAPTED		Producer's Accuracy %
Linear	2m	71.84
	1m	69.83
	0.5m	69.91
Nearest Neighbour	2m	71.17
	1m	73.96
	0.5m	66.13
Cubic	2m	61.54
	1m	70.00
	0.5m	67.46
		65.55

Table J.3: Producer's Accuracy Results for IRI Algorithm

DEFAULT		Producer's Accuracy %
Linear	2m	88.78%
	1m	93.18%
	0.5m	86.81%
Nearest Neighbour	2m	79.80%
	1m	79.59%
	0.5m	72.57%
Cubic	2m	63.57%
	1m	66.67%
	0.5m	75.45%
Raw Data	1m	75.51%

ADAPTED		Producer's Accuracy %
Linear	2m	87.01%
	1m	75.70%
	0.5m	76.29%
Nearest Neighbour	2m	75.76%
	1m	68.87%
	0.5m	66.09%
Cubic	2m	53.41%
	1m	53.49%
	0.5m	50.00%
Raw Data	1m	63.39%

Appendix K

t-Test Results

ATM Default		DATASET		
		LINEAR		
		0.50	1M	2M
	MEAN	1.90	0.82	0.44
	STD DEV	9.97	7.14	6.14
	N	4004001.00	1002001.00	1002001.00
	t-test	381.70	114.66	71.21
	hypothesis	*	*	*
		CUBIC		
		0.50	1M	2M
	MEAN	3.66	1.04	8.55
	STD DEV	12.52	7.19	16.23
	N	4004001.00	1002001.00	1002001.00
	t-test	585.44	145.31	527.45
	hypothesis	*	*	*
		NN		
		0.50	1M	2M
	MEAN	1.18	0.70	0.53
	STD DEV	7.90	6.43	5.99
	N	4004001.00	1002001.00	1002001.00
	t-test	297.87	109.15	88.34
	hypothesis	*	*	*
		RAW DATA		
	MEAN	0.61		
	STD DEV	6.21		
	N	1002001.00		
	t-test	98.79		
	hypothesis	*		

ATM Adapted		DATASET		
		LINEAR		
		0.50	1M	2M
	MEAN	3.08	2.81	2.58
	STD DEV	11.80	11.00	10.58
	N	4004001.00	1002001.00	1002001.00
	t-test	521.65	255.39	243.58
	hypothesis	*	*	*
		CUBIC		
		0.50	1M	2M
	MEAN	15.68	7.69	2.55
	STD DEV	15.86	17.32	10.50
	N	4512255.00	1002001.00	1002001.00
	t-test	2100.36	444.38	243.03
	hypothesis	*	*	*
		NN		
		0.50	1M	2M
	MEAN	10.38	2.32	1.90
	STD DEV	15.05	10.56	9.57
	N	6819408.00	1002001.00	1002001.00
	t-test	1800.97	219.74	198.99
	hypothesis	*	*	*
		RAW DATA		
	MEAN	2.25		
	STD DEV	10.42		
	N	1002001.00		
	t-test	216.39		
	hypothesis	*		

IRI Default		DATASET		
		LINEAR		
		0.50	1M	2M
	MEAN	-0.17	-0.06	0.03
	STD DEV	2.18	2.15	2.92
	N	4006002.00	1003002.00	4006002.00
	t-test	-153.40	-29.39	21.97
	hypothesis	*	*	*
		CUBIC		
		0.50	1M	2M
	MEAN	0.46	0.34	0.28
	STD DEV	2.31	2.24	2.93
	N	4006002.00	4006002.00	4006002.00
	t-test	399.78	300.00	193.74
	hypothesis	*	*	*
		NN		
		0.50	1M	2M
	MEAN	-0.07	-0.04	-0.01
	STD DEV	0.70	0.69	0.66
	N	4008004.00	4008004.00	4008004.00
	t-test	-205.92	-110.74	-24.45
	hypothesis	*	*	*
		RAW DATA		
	MEAN	0.004		
	STD DEV	2.129		
	N	4006002.00		
	t-test	3.76		
	hypothesis	*		

IRI Adapted	DATASET		
	LINEAR		
	0.50	1M	2M
MEAN	-0.23	-0.10	0.00
STD DEV	2.18	2.18	2.94
N	4006002.00	4006002.00	4006002.00
t-test	-209.43	-91.06	1.36
hypothesis	*	*	*
	CUBIC		
	0.50	1M	2M
MEAN	0.46	0.33	0.25
STD DEV	2.32	2.26	2.94
N	4006002.00	4006002.00	4006002.00
t-test	394.26	293.01	172.74
hypothesis	*	*	*
	NN		
	0.50	1M	2M
MEAN	-0.11	-0.06	-0.03
STD DEV	0.78	0.75	0.73
N	4008004.00	4008004.00	4008004.00
t-test	-269.85	-161.02	-82.05
hypothesis	*	*	*
	RAW DATA		
MEAN	-0.02		
STD DEV	2.15		
N	4006002.00		
t-test	-17.66		
hypothesis	*		

Geometric Default	DATASET		
	LINEAR		
	0.50	1M	2M
MEAN	-0.43	0.02	0.39
STD DEV	2.31	1.87	2.92
N	4006002.00	1002001.00	1002001.00
t-test	-375.66	9.64	133.19
hypothesis	*	*	*
	CUBIC		
	0.50	1M	2M
MEAN	-0.03	0.35	0.77
STD DEV	2.55	2.70	3.29
N	4006002.00	1002001.00	1002001.00
t-test	-19.63	129.95	232.85
hypothesis	*	*	*
	NN		
	0.50	1M	2M
MEAN	-0.33	0.07	0.30
STD DEV	1.86	2.16	0.92
N	4004001.00	1002001.00	1002001.00
t-test	-356.40	32.84	328.21
hypothesis	*	*	*

Geometric Adapted	DATASET		
	LINEAR		
	0.50	1M	2M
MEAN	-0.53	0.04	0.52
STD DEV	2.52	2.01	3.19
N	4006002.00	1002001.00	1002001.00
t-test	-424.97	17.47	161.60
hypothesis	*	*	*
	CUBIC		
	0.50	1M	2M
MEAN	-0.09	0.31	0.66
STD DEV	2.65	2.13	3.08
N	4006002.00	1002001.00	1002001.00
t-test	-65.81	144.34	215.80
hypothesis	*	*	*
	NN		
	0.50	1M	2M
MEAN	-0.38	0.01	0.66
STD DEV	2.16	1.08	3.08
N	4004001.00	1002001.00	1002001.00
t-test	-348.97	7.39	215.80
hypothesis	*	*	*

Appendix L

Workings for Research Questions 1-4: Chapter 8

L.1 Q1 Are there significant differences between the classification accuracies for different filtering methods?

IRI DEFAULT vs ATM adapted

	0.5m	1m	2m	Original point density
Linear	1.0171	3.2303	3.4806	
Cubic	4.6704	0.23247	3.0681	
Nearest Neighbour	0.88545	0.58604	0.73831	
Raw Data				0.73831

Hypotheses matrix

	0.5m	1m	2m	Original point density
Linear	-	*	*	
Cubic	*	-	*	
Nearest Neighbour	-	-	-	
Raw Data				-

IRI DEFAULT vs ATM DEFAULT

	0.5m	1m	2m	Original point density
Linear	1.25	3.60	2.63	
Cubic	0.07	2.99	2.32	
Nearest Neighbour	0.72	0.81	0.20	
Raw Data				0.20

Hypotheses matrix

	0.5m	1m	2m	Original point density
Linear	-	*	*	
Cubic	-	*	*	
Nearest Neighbour	-	-	-	
Raw Data				-

IRI DEFAULT vs Geometric default

	0.5m	1m	2m
Linear	2.96	5.17	4.34
Cubic	1.16	1.66	2.40
Nearest Neighbour	0.70	1.21	1.11

Hypotheses matrix

	0.5m	1m	2m
Linear	*	*	*
Cubic	-	-	*
Nearest Neighbour	-	-	-

IRI DEFAULT vs Geometric adapted

	0.5m	1m	2m
Linear	1.48	2.55	1.44
Cubic	1.49	4.42	4.45
Nearest Neighbour	4.36	2.74	2.94

Hypotheses matrix

	0.5m	1m	2m
Linear	-	*	-
Cubic	-	*	*
Nearest Neighbour	*	*	*

IRI adapted vs ATM DEFAULT

	0.5m	1m	2m	Original point density
Linear	1.57	0.89	1.21	
Cubic	6.54	7.11	5.15	
Nearest Neighbour	3.18	3.24	0.92	
Raw Data				0.92

Hypotheses matrix

	0.5m	1m	2m	Original point density
Linear	-	-	-	
Cubic	*	*	*	
Nearest Neighbour	*	*	-	
Raw Data				-

IRI adapted vs ATM DEFAULT

	0.5m	1m	2m	Original point density
Linear	1.57	0.89	1.21	
Cubic	6.54	7.11	5.15	
Nearest Neighbour	3.18	3.24	0.92	
Raw Data				0.92

Hypotheses matrix

	0.5m	1m	2m	Original point density
Linear	-	-	-	
Cubic	*	*	*	
Nearest Neighbour	*	*	-	
Raw Data				-

IRI adapted vs ATM adapted

	0.5m	1m	2m	Original point density
Linear	1.82	1.40	2.05	
Cubic	1.88	4.13	5.89	
Nearest Neighbour	3.34	3.05	1.45	
Raw Data				1.45

Hypotheses matrix

	0.5m	1m	2m	Original point density
Linear	-	-	*	
Cubic	-	*	*	
Nearest Neighbour	*	*	-	
Raw Data				-

IRI adapted vs Geometric default

	0.5m	1m	2m
Linear	1.85	3.73	3.59
Cubic	3.01	3.07	2.28
Nearest Neighbour	4.00	1.71	4.48

Hypotheses matrix

	0.5m	1m	2m
Linear	-	*	*
Cubic	*	*	*
Nearest Neighbour	*	-	*

IRI adapted vs Geometric adapted

	0.5m	1m	2m
Linear	1.37	2.24	0.02
Cubic	7.98	8.55	7.18
Nearest Neighbour	6.76	5.17	3.63

Hypotheses matrix

	0.5m	1m	2m
Linear	-	*	-
Cubic	*	*	*
Nearest Neighbour	*	*	*

ATM DEFAULT VS Geometric default

	0.5m	1m	2m
Linear	2.0736	2.2264	2.1052
Cubic	1.4046	1.5389	0.043112
Nearest Neighbour	0.019807	0.42839	1.513

Hypotheses matrix

	0.5m	1m	2m
Linear	*	*	*
Cubic	-	-	-
Nearest Neighbour	-	-	-
Raw Data			

ATM default VS Geometric adapted

	0.5m	1m	2m
Linear	0.26	1.49	1.46
Cubic	1.58	1.47	2.29
Nearest Neighbour	3.98	2.10	3.00
Raw Data			

Hypotheses matrix

	0.5m	1m	2m
Linear	-	-	-
Cubic	-	-	*
Nearest Neighbour	*	*	*
Raw Data			

ATM adapted VS Geometric default

	0.5m	1m	2m
Linear	2.3605	2.8231	1.0914
Cubic	4.2106	2.1144	0.75169
Nearest Neighbour	0.20722	0.70283	2.1237
Raw Data			

Hypotheses matrix

	0.5m	1m	2m
Linear	*	*	-
Cubic	*	*	-
Nearest Neighbour	-	-	*
Raw Data			

ATM adapted VS Geometric adapted

	0.5m	1m	2m
Linear	0.55	0.96	2.51
Cubic	7.05	5.03	1.53
Nearest Neighbour	3.80	2.39	2.42
Raw Data			

Hypotheses matrix

	0.5m	1m	2m
Linear	-	-	*
Cubic	*	*	-
Nearest Neighbour	*	*	*
Raw Data			

L.2 Q2 Is there a significant difference in accuracy caused by filter method parameter changes?

Showing the Calc Z for the comparison of the two filtering parameters used for the Adaptive TIN Method

ATM DEFAULT vs ATM MINE

	0.5m	1m	2m	Original point density
Linear	0.37	0.71	0.18	
Cubic	2.39	1.85	2.87	
Nearest Neighbour	0.50	1.53	1.06	
Raw Data				1.06

Showing the Hypotheses tests based on comparison of Calc-z and Tab-z for the two parameter settings for the Adaptive TIN Method. A value of - that Calc-z is not greater than Tab-z, and so the null hypothesis is not rejected, indicating that there is no significant difference between the methods. A value of * indicates the converse, and so the null hypothesis must be rejected in favour of the alternative hypothesis which is that there is a significant difference between the two methods.

Hypotheses matrix

	0.5m	1m	2m
Original point density			
Linear	-	-	-
Cubic	*	-	*
Nearest Neighbour	-	-	-
Raw Data			-

L.3 Q3 Is there a significant difference in the classification accuracy between the gridded datasets and the raw data?

Algorithm	Data Structuring Method	Grid space	Z-test result	Hypothesis result
IRI Default	Bilinear	0.5	1.92	-
		1	3.99	*
		2	2.92	*
	Bicubic	0.5	0.18	-
		1	1.73	-
		2	2.61	*
	NN	0.5	0.43	-
		1	0.55	-
		2	0.52	-
IRI Adapted	Bilinear	0.5	2.10	*
		1	2.63	*
		2	4.21	*
	Bicubic	0.5	3.13	*
		1	2.66	*
		2	2.25	*
	NN	0.5	0.30	-
		1	1.17	-
		2	2.63	*
ATM Default	Bilinear	0.5	1.53	-
		1	1.33	-
		2	1.14	-
	Bicubic	0.5	0.38	-
		1	1.66	-
		2	0.22	-
	NN	0.5	0.81	-
		1	2.14	*
		2	1.43	-
ATM Adapted	Bilinear	0.5	0.79	-
		1	0.87	-
		2	0.93	-
	Bicubic	0.5	6.16	*
		1	2.97	*
		2	0.49	-
	NN	0.5	0.06	-
		1	0.87	-
		2	1.01	-

L.4 Q4 Are there significant differences in in the classification accuracies of the different gridded datasets?

ATM default method	Datasets compared	0.5m	1m	2m
	Bilinear/bicubic	1.15	0.35	1.35
Z-test	Bilinear/nearest neighbour	0.75	0.81	0.28
	Bicubic/nearest neighbour	0.42	0.46	1.64
	Datasets compared	**0.5m**	**1m**	**2m**
	Bilinear/bicubic	-	-	-
Hypothesis Result	Bilinear/nearest neighbour	-	-	-
	Bicubic/nearest neighbour	-	-	-

ATM adapted method	Datasets compared	0.5m	1m	2m
	Bilinear/bicubic	6.82	3.82	0.45
Z-test	Bilinear/nearest neighbour	0.85	0.00	1.93
	Bicubic/nearest neighbour	6.13	3.82	1.50
	Datasets compared	**0.5m**	**1m**	**2m**
	Bilinear/bicubic	*	*	-
Hypothesis Result	Bilinear/nearest neighbour	-	-	-
	Bicubic/nearest neighbour	*	*	-

IRI default method	Datasets compared	0.5m	1m	2m
	Bilinear/bicubic	2.12	5.49	5.47
Z-test	Bilinear/nearest neighbour	2.34	3.52	2.42
	Bicubic/nearest neighbour	0.25	2.29	3.14
	Datasets compared	**0.5m**	**1m**	**2m**
	Bilinear/bicubic	*	*	*
Hypothesis Result	Bilinear/nearest neighbour	*	*	*
	Bicubic/nearest neighbour	-	*	*

IRI adapted method	Datasets compared	0.5m	1m	2m
	Bilinear/bicubic	5.08	5.25	6.29
Z-test	Bilinear/nearest neighbour	1.84	1.45	1.68
	Bicubic/nearest neighbour	3.48	3.79	4.82
	Datasets compared	**0.5m**	**1m**	**2m**
	Bilinear/bicubic	*	*	*
Hypothesis Result	Bilinear/nearest neighbour	-	-	-
	Bicubic/nearest neighbour	*	*	*

Geometric default method	Datasets compared	0.5m	1m	2m
	Bilinear/bicubic	0.48	1.08	0.82
Z-test	Bilinear/nearest neighbour	1.34	3.52	0.90
	Bicubic/nearest neighbour	1.83	2.49	0.09
	Datasets compared	**0.5m**	**1m**	**2m**
	Bilinear/bicubic	-	-	-
Hypothesis Result	Bilinear/nearest neighbour	-	*	-
	Bicubic/nearest neighbour	-	*	-

Geometric adapted method	Datasets compared	0.5m	1m	2m
	Bilinear/bicubic	0.48	1.08	0.82
Z-test	Bilinear/nearest neighbour	1.34	3.52	0.90
	Bicubic/nearest neighbour	1.83	2.49	0.09
	Datasets compared	**0.5m**	**1m**	**2m**
	Bilinear/bicubic	-	-	-
Hypothesis Result	Bilinear/nearest neighbour	-	*	-
	Bicubic/nearest neighbour	-	*	-

Appendix M

Workings for Research Question 5: Chapter 8

Aim of investigation: Testing for significant difference in classification accuracies when using datasets at different spatial resolutions

Table M.1: Results for default IRI algorithm

Dataset	Comparison	Z-test result	Hypothesis result
Bilinear	0.5m vs 1m	2.27	*
	0.5m vs 2m	1.00	-
	1m vs 2m	1.38	-
Bicubic	0.5m vs 1m	1.57	-
	0.5m vs 2m	2.47	*
	1m vs 2m	0.89	-
Nearest	0.5m vs 1m	0.98	-
Neighbour	0.5m vs 2m	0.96	-
	1m vs 2m	0.02	-

Table M.2: Results for adapted IRI algorithm

Dataset	Comparison	Z-test result	Hypothesis result
Bilinear	0.5m vs 1m	0.47	-
	0.5m vs 2m	2.16	*
	1m vs 2m	1.74	-
Bicubic	0.5m vs 1m	0.52	-
	0.5m vs 2m	0.96	-
	1m vs 2m	0.44	-
Nearest	0.5m vs 1m	0.89	-
Neighbour	0.5m vs 2m	2.37	*
	1m vs 2m	1.46	-

Table M.3: Results for adapted ATM algorithm

Dataset	Comparison	Z-test result	Hypothesis result
Bilinear	0.5m vs 1m	0.20	-
	0.5m vs 2m	0.38	-
	1m vs 2m	0.18	-
Bicubic	0.5m vs 1m	1.29	-
	0.5m vs 2m	0.59	-
	1m vs 2m	1.87	-
Nearest	0.5m vs 1m	1.37	-
Neighbour	0.5m vs 2m	0.64	-
	1m vs 2m	0.72	-

Table M.4: Results for default Geometric algorithm

Dataset	Comparison	Z-test result	Hypothesis result
Bilinear	0.5m vs 1m	0.37	-
	0.5m vs 2m	0.44	-
	1m vs 2m	0.07	-
Bicubic	0.5m vs 1m	1.20	-
	0.5m vs 2m	0.87	-
	1m vs 2m	0.33	-
Nearest	0.5m vs 1m	1.84	-
Neighbour	0.5m vs 2m	0.87	-
	1m vs 2m	2.69	*

Table M.5: Results for adapted Geometric algorithm

Dataset	Comparison	Z-test result	Hypothesis result
Bilinear	0.5m vs 1m	1.58	-
	0.5m vs 2m	1.36	-
	1m vs 2m	0.22	-
Bicubic	0.5m vs 1m	1.25	-
	0.5m vs 2m	0.13	-
	1m vs 2m	1.11	-
Nearest	0.5m vs 1m	0.53	-
Neighbour	0.5m vs 2m	0.35	-
	1m vs 2m	0.19	-

Appendix N

Results and Workings for Research Question 2: Chapter 9

This appendix presents the results for the following research question: Is there a significant difference caused by changing parameter settings for segmentation methods? How sensitive are the methods?

N.1 Comparison of default and adapted settings for Geometric algorithm

Table N.1: $Z - test$ and hypotheses results for Geometric algorithm

Dataset	Resolution	Class	Z-TEST	Hypothesis
Linear	0.5	Ground	0.45	-
		Building	0.78	-
		Vegetation	0.61	-
	1m	Ground	1.60	-
		Building	0.17	-
		Vegetation	0.59	-
	2m	Ground	0.79	-
		Building	0.59	-
		Vegetation	0.35	-
Cubic	0.5	Ground	0.17	-
		Building	1.09	-
		Vegetation	0.33	-
	1m	Ground	1.26	-
		Building	0.56	-
		Vegetation	0.11	-
	2m	Ground	0.82	-
		Building	0.65	-
		Vegetation	0.10	-
NN	0.5	Ground	0.00	-
		Building	0.83	-
		Vegetation	2.81	*
	1m	Ground	2.24	*
		Building	0.00	-
		Vegetation	0.12	-
	2m	Ground	1.49	-
		Building	0.55	-
		Vegetation	1.41	-

N.2 Comparison of default and adapted settings for Geometric algorithm

Table N.2: $Z - test$ and hypotheses results for ATM algorithm

Dataset	Resolution	Class	Z-TEST	Hypothesis
Linear	0.5	Ground	0.17	-
		Building	0.32	-
		Vegetation	0.00	-
	1m	Ground	0.71	-
		Building	0.66	-
		Vegetation	0.40	-
	2m	Ground	0.33	-
		Building	0.93	-
		Vegetation	0.00	-
Cubic	0.5	Ground	0.39	-
		Building	0.97	-
		Vegetation	0.00	-
	1m	Ground	3.20	*
		Building	0.94	-
		Vegetation	1.59	-
	2m	Ground	0.43	-
		Building	1.40	-
		Vegetation	0.88	-
NN	0.5	Ground	0.22	-
		Building	1.87	-
		Vegetation	0.65	-
	1m	Ground	1.66	-
		Building	1.72	-
		Vegetation	0.36	-
	2m	Ground	1.11	-
		Building	1.77	-
		Vegetation	0.57	-

Appendix O

Results and Workings for Research Question 4: Chapter 9

This appendix presents the results to the following research questions: Are there significant differences in classification accuracy for different gridded datasets?

Table O.1: Results for Geometric algorithm using Default Settings

Geometric Default

0.5m

Z-TEST

	ground	buildings	vegetation
linear vs cubic	0.26	0.00	0.27
linear vs nearest	1.94	2.35	0.38
cubic vs nearest	0.92	0.83	1.77

HYPOTHESES

	ground	buildings	vegetation
linear vs cubic	-	-	-
linear vs nearest	-	*	-
cubic vs nearest	-	-	-

1m

Z-TEST

	ground	buildings	vegetation
linear vs cubic	0.08	0.33	0.25
linear vs nearest	0.85	1.20	0.34
cubic vs nearest	0.28	0.14	0.14

HYPOTHESES

	ground	buildings	vegetation
linear vs cubic	-	-	-
linear vs nearest	-	-	-
cubic vs nearest	-	-	-

2m

Z-TEST

	ground	buildings	vegetation
linear vs cubic	1.04	0.18	0.86
linear vs nearest	0.22	0.63	0.85
cubic vs nearest	0.47	3.21	2.67

HYPOTHESES

	ground	buildings	vegetation
linear vs cubic	-	-	-
linear vs nearest	-	-	-
cubic vs nearest	-	*	*

Table O.2: Results for Geometric algorithm using Adapted Settings

Geometric Adapted

0.5m

Z-TEST

	ground	buildings	vegetation
linear vs cubic	1.42	0.35	1.07
linear vs nearest	0.47	3.08	2.53
cubic vs nearest	1.27	0.65	1.95

HYPOTHESES

	ground	buildings	vegetation
linear vs cubic	-	-	-
linear vs nearest	-	*	*
cubic vs nearest	-	-	-

1m

Z-TEST

	ground	buildings	vegetation
linear vs cubic	0.70	0.53	0.17
linear vs nearest	0.33	0.76	0.43
cubic vs nearest	0.54	0.13	0.41

HYPOTHESES

	ground	buildings	vegetation
linear vs cubic	-	-	-
linear vs nearest	-	-	-
cubic vs nearest	-	-	-

2m

Z-TEST

	ground	buildings	vegetation
linear vs cubic	0.17	0.09	0.09
linear vs nearest	0.22	0.86	0.63
cubic vs nearest	2.42	1.06	1.29

HYPOTHESES

	ground	buildings	vegetation
linear vs cubic	-	-	-
linear vs nearest	-	-	-
cubic vs nearest	*	-	-

Table O.3: Results for ATM algorithm using Default Settings

ATM Default

0.5m

Z-TEST	ground	buildings	vegetation
linear vs cubic	0.49	0.25	0.24
linear vs nearest	1.45	1.01	0.43
cubic vs nearest	1.35	0.34	1.00

HYPOTHESES	ground	buildings	vegetation
linear vs cubic	-	-	-
linear vs nearest	-	-	-
cubic vs nearest	-	-	-

1m

Z-TEST	ground	buildings	vegetation
linear vs cubic	0.33	0.85	0.51
linear vs nearest	0.30	0.98	0.68
cubic vs nearest	1.24	1.62	0.37

HYPOTHESES	ground	buildings	vegetation
linear vs cubic	-	-	-
linear vs nearest	-	-	-
cubic vs nearest	-	-	-

2m

Z-TEST	ground	buildings	vegetation
linear vs cubic	1.15	0.42	0.73
linear vs nearest	1.37	0.77	0.59
cubic vs nearest	0.65	1.13	0.48

HYPOTHESES	ground	buildings	vegetation
linear vs cubic	-	-	-
linear vs nearest	-	-	-
cubic vs nearest	-	-	-

Table O.4: Results for ATM algorithm using Default Settings

ATM Adapted

0.5m

Z-TEST	ground	buildings	vegetation
linear vs cubic	1.37	0.08	1.28
linear vs nearest	4.46	1.05	3.20
cubic vs nearest	3.40	0.53	4.04

HYPOTHESES	ground	buildings	vegetation
linear vs cubic	-	-	-
linear vs nearest	*	-	*
cubic vs nearest	*	-	*

1m

Z-TEST	ground	buildings	vegetation
linear vs cubic	0.00	0.24	0.24
linear vs nearest	0.33	0.55	0.22
cubic vs nearest	2.37	1.52	4.11

HYPOTHESES	ground	buildings	vegetation
linear vs cubic	-	-	-
linear vs nearest	-	-	-
cubic vs nearest	*	-	*

2m

Z-TEST	ground	buildings	vegetation
linear vs cubic	0.75	0.42	0.33
linear vs nearest	0.21	0.10	0.10
cubic vs nearest	0.36	1.05	0.69

HYPOTHESES	ground	buildings	vegetation
linear vs cubic	-	-	-
linear vs nearest	-	-	-
cubic vs nearest	-	-	-

Appendix P

Results and Workings for Research Question 6: Chapter 9

This appendix presents the results used to answer the following research question: are there differences between the classification accuracies for the two above ground classes?

Table P.1: Comparison of Above Ground Classification Accuracies for Geometric Algorithm with Default Settings

		z-test	H
Linear	0.5m	0.71	-
	1m	2.46	*
	2m	2.40	*
Cubic	0.5m	2.19	*
	1m	3.18	*
	2m	2.58	*
Nearest	0.5m	1.03	-
	1m	3.73	*
	2m	6.56	*

Table P.2: Comparison of Above Ground Classification Accuracies for Geometric Algorithm with Adapted Settings

		z-test	H
Linear	0.5m	0.52	-
	1m	2.92	*
	2m	4.75	*
Cubic	0.5m	1.16	-
	1m	3.85	*
	2m	2.60	*
Nearest	0.5m	3.31	*
	1m	3.95	*
	2m	4.65	*

Table P.3: Comparison of Above Ground Classification Accuracies for ATM Algorithm with Default Settings

		z-test	H
Linear	0.5m	3.68	*
	1m	5.14	*
	2m	3.96	*
Cubic	0.5m	3.31	*
	1m	4.17	*
	2m	1.77	-
Nearest	0.5m	2.85	*
	1m	4.61	*
	2m	1.99	*
Raw		3.43	*

Table P.4: Comparison of Above Ground Classification Accuracies for ATM Algorithm with Adapted Settings

		z-test	H
Linear	0.5m	3.02	*
	1m	2.11	*
	2m	3.31	*
Cubic	0.5m	1.39	-
	1m	4.85	*
	2m	3.20	*
Nearest	0.5m	1.37	-
	1m	1.36	-
	2m	2.44	*
Raw		2.7766	*

Bibliography

Abeyta, A. M. and J. Franklin (1998). The accuracy of vegetation stand boundaries derived from image segmentation in a desert environment. *Photogrammetric Engineering and Remote Sensing 64*, 59–66.

Agresti, A. (1996). *An Introduction to Categorical Data Analysis*. Wiley, New York.

Ahokas, E., H. Kaartinen, and J. Hyyppa (2003). A quality assessment of airborne laser scanner data. In *Proceedings of ISPRS Working Group III/3 3-D Reconstruction from airborne laserscanner and InSAR Data*, Volume XXXIV, Part 3/W13, Dresden.

Altharthy, A. and J. Bethel (2003). Heuristic filtering and 3d feature extraction from lidar data. In *Proceedings of ISPRS Commission III/WG3 3D Reconstruction from Laser Scanning and InSAR*, Dresden, Germany.

Amidror, I. (2002). Scattered data interpolation methods for electronic imaging systems: a survey. *Journal of Electronic Engineering 11*(2), 157–176.

Anselin, L., Y. W. Kim, and I. Syabri (2003). Web-based analytical tools for the exploration of spatial data. Unpublished manuscript.

Aplin, P., P. M. Atkinson, and P. J. Curran (1997). Fine resolution spatial satellite sensors for the next decade. *International Journal of Remote Sensing 18*, 3873–3882.

Appleton, K., A. Lovett, T. Dockerty, and G. Sünnenberg (2004). Representing uncertainty in visualisations of future landscapes. In *ISPRS 20th Congress Proceedings*. DVD collection.

Armstrong, M. (1998). *Basic Linear Geostatistics.* Springer.

ASPRS (1989). American Society for Photogrammetry and Remote Sensing (ASPRS) interim accuracy standards for large scale line maps. *Photogrammetric Engineering and Remote Sensing 55*, 1038–1040.

Atkinson, P. M. and P. J. Curran (1997). Choosing and appropriate resolution for remote sensing investigations. *Photogrammetric Engineering and Remote Sensing 63*, 1345–51.

Axelsson, P. (1999). Processing of laser scanner data - algorithms and applications. *ISPRS Journal of Photogrammetry and Remote Sensing 54*, 138–147.

Axelsson, P. (2000). DEM generation from laser scanner data using adaptive TIN models. In *Proceedings of IAPRS*, Volume 33, pp. 110–117.

Axelsson, P. (2001). Ground estimation of laser data using adaptive TIN-models. In *Proceedings of OEEPE workshop on airborne laserscanning and interferometric SAR for detailed elevation models*, Number 40: Official Publication, Stockholm, Sweden, pp. 185–208.

Bailey, T. C. and A. C. Gatrell (1995). *Interactive Spatial Data Analysis.* Longman, England.

Baker, T. L. (1994). *Doing Social Research.* McGraw Hill Inc, New York. 2nd Edition.

Balce, A. E. (1987). Determination of optimal sampling interval in grid digital elevation models (dem) data aquisition. *Photogrammetric Engineering and Remote Sensing 53*, 323–330.

Baltsavias, E. P. (1999). A comparison between photogrammetry and laser scanning. *ISPRS Journal of Photogrammetry and Remote Sensing 54*, 83–94.

Baltsavias, E. P. and A. Gruen (2003). Resolution convergence: A comparison of aerial photos, lidar, and ikonos for modelling cities. In V. Mesev (Ed.), *Remotely Sensed Cities*, pp. 47–82. Taylor and Francis, London.

Balz, T. and N. Haala (2003). SAR-based 3D-reconstruction of complex urban

environments. In *Proceedings of ISPRS commission III workshop: 3D Reconstruction from InSAR and LiDAR*, Dresden, Germany.

Barnsley, M. J. and S. L. Barr (1996). Inferring urban land use from satellite sensor images using kernel-based spatial reclassification. *Photogrammetric Engineering and Remote Sensing 62*(8), 949–958.

Barnsley, M. J. and S. L. Barr (2000). Monitoring urban land use by earth observation. *Surveys in Geophysics 21*, 269–289.

Barnsley, M. J., A. M. Steel, and S. L. Barr (2003). Determining urban land use through an analysis of the spatial composition of buildings identified in lidar and multispectral image data. In V. Mesev (Ed.), *Remotely Sensed Cities*, pp. 84–108. Taylor and Francis, London.

Barr, S. and M. J. Barnsley (1997). A region-based, graph-theoretic data model for the inference of second-order information from remotely-sensed images. *International Journal of Geographical Information Science 11*(6), 555–576.

Barr, S. L. and M. J. Barnsley (1998a). Application of structural pattern-recognition techniques to infer urban land use from Ordnance Survey digital map data. In *Geocomputation98 3rd International Conference on Geocomputation.* CD-Rom collection.

Barr, S. L. and M. J. Barnsley (1998b). A syntactic pattern recognition paradigm for the derivation of second-order thematic information from remotely-sensed images. In P. Atkinson and N. Tate (Eds.), *Advances in Remote Sensing and GIS Analysis*, Chapter 11, pp. 167–184. John Wiley and Sons.

Barr, S. L. and M. J. Barnsley (2004). Quantifying the separability of urban land use via a structural pattern recognition system. *Environment and Planning B 31*, 397–418.

Batty, M. and D. Howes (2001). Predicting temporal patterns in urban development from remote imagery. In J.-P. Donnay, M. J. Barnsley, and P. A. Longley (Eds.), *Remote Sensing and Urban Analysis*, pp. 185–204. Taylor and Francis, London.

Bean, J. E. and C. R. Ferguson (2003). Effective use of the ConnDOT GPS base station. Technical report, University of Connecticut, Connecticut Transportation Institute, Storrs, CT.

Bechini, L., S. Bocchi, and T. Maggiore (2003). Spatial interpolation of soil physical properties for irrigation planning. a simulation study in northern Italy. *European Journal of Agronomy 19*(1), 1–14.

Behan, A. (2000). On the matching accuracy of rasterised scanning laser altimetry data. *International Archives of Photogrammetry and Remote Sensing XXXIII Part 2.* Amsterdam.

Behan, A., H.-G. Maas, and G. Vosselman (2000). Steps towards quality improvement of airborne laser scanner data. In *Proceedings of 26th Annual Conference of the Remote Sensing Society.*

Bellman, C. and M. Shortis (2000). Early stage object recognition using neural networks. *International Archives of Photogrammetry and Remote Sensing Vol XXXIII*(Supplement B3), 20–26.

Brenner, C. (2005). *International Journal of Applied Earth Observation and Geoinformation 6*(3-4), 187–198.

Briese, C., N. Pfeifer, and P. Dorninger (2002). Applications of the robust interpolation for DTM determination. In *International Archives of Photogrammetry and Remote Sensing, Volume XXXIV, Symposium of the ISPRS-Comm. III, Graz, 9 - 13 September 2002.*

Briggs, D. and S. Butterworth (1981). The principles and practice of applied geography. *Applied Geography 1*, 1–8.

Brovelli, M. A., M. Cannata, and U. M. Longoni (2002). Managing and processing lidar data within grass. In *Proceedings of the GRASS Users Conference*, Trento.

Brown, D. and T. Bara (1994). Recognition and reduction of systematic error in elevation and derivative surfaces from 7.5-minute dems. *Photogrammetric Engineering and Remote Sensing 60*(2), 189–194.

Brown, D. G. and J. B. Thaddeus (1994). Recognition and reduction of systematic

error in elevation and derivative surfaces from 7 1/2 minute dems. *Photogrammetric Engineering and Remote Sensing 60*(2), 189–194.

Brunn, A. and U. Weidner (1997). Extracting buildings from digital surface models. *International Archives of Photogrammetry and Remote Sensing 32*(3-4W2).

Buckley, S. J. and H. L. Mitchell (2004, December). Integration, validation and point spacing optimisation of digital elevation models. *The Photogrammetric Record 19*(108), 277–296.

Bunkin, A. F., D. V. Vlasov, and R. A. Garaev (1983). *Soviet Journal of Quantitative Electronics 10*(9).

Burrough, P. A. (1986). *Principles of Geographical Information Systems for Land Resources Assessment.* OUP, Oxford.

Burrough, P. A. and R. A. McDonnell (1998). *Principles of Geographical Information Systems.* Oxford University Press. New York.

Carlisle, B. H. (2000). The highs and lows of digital elevation model error: creating a spatially distributed dem error mode. In R. J. Abrahart and B. H. Carlisle (Eds.), *GeoComputation 2000: Proceedings of the 5th International Conference on GeoComputation*, Manchester, UK.

Carlson, T. (2003). Preface: Applications of remote sensing to urban problems. *Remote Sensing of the Environment 86*, 273–274.

Carmines, E. G. and R. A. Zeller (1979). *Reliability and Validity Assessment.* Sage University Papers. Series: Quantitative Applications in the Social Sciences (Number 07-017) Edited by J L Sullivan.

Carter, J. R. (1988). Digital representations of topographic surfaces. *Photogrammetric Engineering and Remote Sensing 54*(11), 1577–80.

Carter, J. R. (1989). Relative errors identified in usgs gridded dems. In *Proceedings, Auto Carto 9*, Bethesda, Maryland, pp. 255–265.

Carter, J. R. (1992). The effect of data precision on the calculation of slope and aspect using gridded dems. *Cartographica 29*(1), 22–34.

Chorley, R. J. and P. Haggett (1965). Trend surface models in geographical research. *Transactions of the Institute of British Geographers 37*, 47–67.

Clark, W. A. V. and P. L. Hosking (1986). *Statistical Methods for Geographers.* Wiley, New York.

Cobby, D. M., D. C. Mason, and I. J. Davenport (2001). Image processing of airborne scanning laser altimetry data for improved river flood mapping. *ISPRS Journal of Photogrammetry and Remote Sensing 56*, 121–138.

Couloigner, I. and T. Ranchin (2000). Mapping of urban areas: A multiresolution modelling approach for semi-automatic extraction of streets. *Photogrammetric Engineering and Remote Sensing 66*, 867–74.

Cramer, M. (1999). Direct geocoding - is aerial triangulation obsolete? *Photogrammetric Week 99*(47), 59–69.

Crombaghs, M. J. E., B. R, and E. J. D. Min (2000). On the adjustment of overlapping strips of laseraltimeter height data. *ISPRS 33 B3/1*, 230–237.

Dahany, J. (1997). A set of visualisation data needs in urban environmental planning and design for photogrammetric data. In *Proceedings of the Ascona workshop '97: Automatic extraction of man-made objects from aerial and space images*, Monte Verita, Switzerland, pp. 357–365.

Dash, J., E. Steinle, R. P. Singh, and H. P. Bähr (2004). Automatic building extraction from laser scanning data: an input tool for disaster management. *Advances in Space Research 33*, 317–322.

Davis, B. M. (1987). Uses and abuses of cross-validation in geostatistics. *Mathematical Geology 19*(3), 241–248.

Declercq, F. A. N. (1996). Interpolation methods for scattered sample data: Accuracy, spatial patterns, processing time. *Cartography and Geographical Information Systems 23*(3), 128–144.

Delaunay, B. (1934). Sur la sphère vide. *Bulletin de l'Académie Des Sciences de l'Union des Républiques Soviétiques Socialistes, Classes des Sciences Mathématiques et Naturelles Série VII*, 793–800.

Desmet, P. J. J. (1997). Effects of interpolation errors on the analysis of DEMs. *Earth Surface Processes and Landforms 22*, 563–580.

Deutsch, C. V. and A. G. Journel (1992). *GSLIB Geostatistical Software Library and User's Guide*. Oxford University Press. New York.

Diamantopoulos, A. and B. Schlegelmilch (1997). *Taking the Fear Out of Data Analysis*. Dryden Press, London.

Dodge, M., A. Smith, and S. Doyle (1997). Visualising urban environments for planning and design. In *Proceedings of Graphics, Visualization and the Social Sciences workshop*, Loughborough, UK.

Donnay, J.-P., M. J. Barnsley, and P. A. Longley (2003). Introduction. In V. Mesev (Ed.), *Remotely Sensed Cities*, pp. 3–18. Taylor and Francis, London.

Donnay, J.-P. and I. Nadasdi (1992). Usages des données satellitaires urbaines de haute résolution en modélisation urbaine: application à l'agglomération de maastricht. *Acta Geographica Lovaniensia 33*, 159–169.

Dowman, I. and P. Fischer (2003). Evaluation of IfSAR and LiDAR data for flood risk assessment. In *Proceedings of ISPRS Workshop*, Portland, Oregon, USA.

EA (1998). Airborne light detection and rangin feasibility study. Rd technical summary es421, Bristol: National Centre for Environmental Data and Surveillance, Environment Agency.

Eason, G., C. W. Coles, and G. Gettinby (1980). *Mathematics and Statistics for the Bio-Sciences*. John Wiley and Sons.

Edwards, L. D. and E. S. Nelson (2001). Visualizing data uncertainty: A case study using graduated circles maps. *Cartographic Perspectives 38*, 19–36.

Edzer, J. and C. G. Wessling (1998). Gstat: a program for geostatistical modelling, prediction and simulation. *Computers and Geosciences 24*(1), 17–31.

Efron, B. (1982). *The Jackknife, the Bootstrap and Other Resampling Plans*. Society for Industrial and Applied Mathematics.

Efron, B. (1993). *An Introduction to the Bootstrap*. Chapman and Hall. Monographs on Statistics and Applied Probability, 57.

Ehlschlaeger, C. (1998). *The Stochastic Simulation Approach: Tools for representing spatial application uncertainty.* Ph. D. thesis, University of California, Santa Barbara.

Ehlschlaeger, C. R. and M. F. Goodchild (1994). Uncertainty in spatial data: defining, visualising, and managing data errors. In *Proceedings of GIS/LIS*, Phoenix, Arizona, pp. 246–53.

Eliasson, I. (1992). Infrared thermography and urban temperature patterns. *International Journal of Remote Sensing 13*, 869–880.

Elmqvist, M. (2001). Ground estimation of laser radar data using active shape models. In *Proceedings of OEEPE Workshop on Airborne Laserscanning and Interferometric SAR for Detailed Digital Elevation Models*, Royal Institute of Technology Department of Geodesy and Photogrammetry, Stockholm, Sweden.

Elmqvist, M., M. Jungert, F. Lantz, A. Persson, and U. Södermann (2001). Terrain modelling and analysis using laser scanner data. *International Archives of Photogrammetry and Remote Sensing, Annapolis, Maryland,* USA *XXXIV-3/W4.*

Fan, H. and D. J. Sailor (2005). Modeling the impacts of anthropogenic heating on the urban climate of philadelphia: A comparison of implementations in two pbl schemes. *Atmospheric Environment 39*(1), 73–84.

Felicísimo, A. M. (1994). Parametric statistical method for error detection in digital elevation models. *ISPRS Journal of Photogrammetry and Remote Sensing 49*(4), 29–33.

Filin, S. (2002, September 9-11). Surface clustering from airborne laser scanning data. *International Archives of Photogrammetry and Remote Sensing XXXIV*(Part 3A), 119–124. Graz, Austria.

Fiocco, G. and L. D. Smullin (1963). Detection of scattering layers in the upper atmosphere (60-140km) by optical radar. *Nature 199*, 1275–1276.

Fischer, P. and I. Dowman (2002). Validation of ifsar digital elevation models from intermap technologies. Technical report, Norwich Union Insurance.

Fisher, P. F. (1988). Improved modelling of elevation error with geostatistics. *Geoinformatica 2*, 215–33.

Fisher, P. F. (1990). Simulation of error in Digital Elevation Models. In *Papers and Proceedings of the 13th Applied Geography Conference*, Volume 13, pp. 37–43.

Fisher, P. F. (1991). First experiments in viewshed uncertainty: The accuracy of the viewshed area. *Photogrammetric Engineering and Remote Sensing 57*(10), 1321–1327.

Fisher, P. F. (1992). First experiments in viewshed uncertainty: Simulating fuzzy viewshed. *Photogrammetric Engineering and Remote Sensing 58*(3), 345–352.

Flood, M. (2001). Lidar activities and research priorities in the commercial sector. In *Proceedings of IAPRS WG IV/3*, Volume XXXIV, pp. 678–684.

Flood, M. and B. Gutelius (1997). Commercial implications of topographic terrain mapping using airborne laser radar. *Photogrammetric Engineering 63*, 327–329; 363–366.

Florinsky, I. V. (2002). Errors of signal processing in digital terrain modelling. *Internation Journal of Geographical Information Science 16*(5), 475–501.

Foody, G. (2004). Thematic map comparison: Evaluating the statistical significance of differences in classification accuracy. *Photogrammetric Engineering and Remote Sensing 70*(5), 627–633.

Foody, G. M. (2000). Estimation of sub-pixel land cover composition in the presence of untrained classes. *Computers and Geosciences 26*, 469–478.

Foody, G. M. and P. M. Atkinson (2002). *Uncertainty in Remote Sensing and GIS*. Wiley. Chichester.

Foody, G. M. and D. S. Boyd (1999). Fuzzy mapping of tropical land cover along an environmental gradient from remotely sensed data with an artificial neural network. *Journal of Goegraphical Systems 1*, 23–35.

Forster, B. C. (1983). Some urban measurements from landsat data. *Photogrammetric Engineering and Remote Sensing 49*, 1693–1707.

Frederiksen, P., O. Jacobi, and K. Kubik (1998). A review of current trends in terrain modelling. *I.T.C. Journal 2*, 101–106.

Fuchs, C. (1996). OEEPE study on 3D city models. In *Report of the Institute of Photogrammetry*, University of Bonn.

Gamba, P., B. Houshmand, and M. Saccani (2000). Detection and extraction of buildings from interferometric SAR data. *IEEE Transactions on Geoscience and Remote Sensing 38*(1), 611–618.

Gardner, T. W., K. C. Sasowsky, and R. L. Day (1990). Automated extraction of geomorphometric properties from digital elevation data. *Zeitschrift fur geomorphologie 80*, 57–68.

Geoghegan, J., J. L. Pritchard, Y. Ogneva-Himmelberger, R. R. Chowdhury, S. Sanderson, and B. L. Turner (1998). Socialising the pixel" and "pixelising the social" in land use and land-cover change. In D. Liverman, E. F. Moran, R. R. Rindfuss, and P. C. Stern (Eds.), *People and pixels: linking remote sensing and social science*, pp. 51–69. National Academy Press, Washington DC.

Gerke, M., C. Heipke, and B.-M. Straub (2001). Building extraction form aerial imagery using a generic scene models and invariant geometric moments. In *IEEE/ISPRS Joint Workshop on Remote Sensing and Data Fusion over Urban Areas.*

Gillies, R. R., J. B. Box, J. Sysanzik, and E. J. Rodemaker (2003). Effects of urbanisation on the aquatic fauna of the Line Creek watershed, Atlanta - a satellite perspective. *Remote Sensing of Environment 86*, 411–422.

Gomarasca, M. A., P. A. Brivio, F. Pagnoni, and A. Galli (1993). One century of land use changes in the metropolitan area of Milan (Italy). *International Journal of Remote Sensing 14*, 211–223.

Goodchild, M. F. and S. Gopal (1989). *The Accuracy of Spatial Databases.* Taylor and Francis. London.

Grey, W. M. F. and A. J. Luckman (1999). Using SAR and interferometric phase coherence to detect urban change. In *RSS'99 - Earth Observation: from Data to*

Information, Cardiff, UK. Remote Sensing Society International Symposium.

Gruen, A. and H. Dan (1997). Tobago - a topology builder for the automated generation of building models. In *Proceedings of the Ascona Workshop '97: Automatic extraction of man-made objects from aerial and space images.*

Guptill, S. C. and J. Morrison (1995). *The Elements of Spatial Data Quality.* Elsevier, Amsterdam.

Guth, P. (1992). Spatial analysis of dem error. In *Proceedings ASPRS/ACSM Annual Meeting*, Washington DC, pp. 187–196.

Haala, N. and C. Brenner (1999). Extraction of buildings and trees in urban environments. *ISPRS Journal of Photogrammetry and Remote Sensing 54*, 130–137.

Haala, N., C. Brenner, and K. Anders (1998). 3D urban gis from laser altimetry and 2D map data. *International Archives of Photogrammetry and Remote Sensing 32*(3/1), 339–346.

Haala, N. and K. Brenner (1997). Generation of 3D city models from airborne laser scanning data. In *Proceedings EARSEL Workshop on LIDAR remote sensing on land and sea*, Tallinn, Estonia.

Hammer, A., D. Heineann, C. Hoyer, R. Kuhlemann, E. Lorenz, R. Müller, and H. G. Beyer (2003). Solar energy assessment using remote sensing technologies. *Remote Sensing of Environment 86*, 423–432.

Hannah, M. J. (1981). Error detection and correction in digital terrain models. *Photogrammetric Engineering and Remote Sensing 47*(1), 63–69.

Harrower, M. (2003). Representing uncertainty: Does it help people make better decisions? (invited white paper). In *UCGIS Workshop: Geospatial Visualization and Knowledge Discovery*, National Conference Center, Landsdowne, V.A.

Haugerud, R. A. and D. J. Harding (2001). Some algorithms for virtual deforestation (VDF) of LIDAR topographic survey data. In *Proceedings of IAPRS*, Volume XXXIV-3/W4, Annapolis, MD, pp. 211–218.

Hengl, T. (2003). Visualisation of uncertainty using the HSI colour model: computations with colours. In *Proceedings of 7th International Conference on GeoComputation*, pp. 8.

Hepner, G. F., B. Houshmand, I. Kulikov, and N. Bryant (1998). Investigation of the integration of AVIRIS and IFSAR for urban analysis. *Photogrammetric Engineering and Remote Sensing 64*(8), 813–820.

Heuvelink, G. B. M. (1998). *Error Propagation in Environmental Modelling.* Taylor and Francis.

Hofmann, A., H.-G. Maas, and A. Streilein (2002). Knowledge-based building detection based on laser scanner data and topographic map information. In *International Archives of Photogrammetry and Remote Sensing*, Volume XXXIV 3A III, Graz, Austria.

Holland, D. A. and P. Marshall (2003). Using high-resolution satellite imagery in a well-mapped country. In *Proc.of ISPRS-EARSeL Joint Work. on High Resolution Mappingfrom Space*, Hanover. On CDROM.

Huising, E. J. and L. M. G. Pereira (1998a). Errors and accuracy estimates of laser altimetry data aquired by various laser scanning systems for topographic applications. *ISPRS Journal of Photogrammetry and Remote Sensing 53*(5), 245–261.

Huising, E. J. and L. M. G. Pereira (1998b). Errors and accuracy estimates of laser data acquired by various laser scanning systems for topographic applications. ISPRS *Journal of Photogrammetry and Remote Sensing 53*(5), 245–261.

Hunter, G. J. and M. F. Goodchild (1997). Modeling the uncertainty of slope and aspect estimates derived from spatial databases. *Geographical Analysis 29*, 35–49.

Itti, L., C. Koch, and E. Niebur (1998). A model of saliency – based visual attention for rapid scene analysis. *IEEE Trans. Pattern Anal. Mach. Intell. (PAMI) 20*(11), 1254–1259.

Jelalian, A. V. (1992). *Laser Radar Systems.* Artexh House.

Kofler, M. and M. Gruber (1997). Toward a 3D GIS database. *GIM 11*(5), 55–59.

Krabhill, W. B., C. W. Wright, R. N. Swift, and E. B. Frederick (1995). Accuracy of airborne laser altimetry over the Greenland ice sheet. *International Journal of Remote Sensing 16*(7), 1211–1222.

Kraus, K. and N. Pfeifer (1998). Determination of terrain models in wooded areas with airborne laser scanner data. ISPRS *Journal of Photogrammetry and Remote Sensing 53*, 193–203.

Kuilenburg, J. V., J. J. D. Gruigter, B. A. Marsman, and J. Bouma (1982). Accuracy of spatial interpolation between point data on soil moisture supply capacity, compared with estimates from mapping units. *Geoderma*, 311–25.

Kumler, M. P. (1994). An intensive comparison of triangulated irregular networks (TINs) and Digital Elevation Models (DEMs). *Cartographica, Monograph 31*(2), 1–99.

Lam, N. S. N. (1983). Spatial interpolation methods: A review. *The American Cartographer 10*(2), 129–149.

Laslett, G. M., A. B. McBratney, P. J. Phal, and M. F. Hutchinson (1987). Comparison of several spatial prediction methods for soil ph. *Journal of Soil Science 38*, 325–41.

Lazzaro, D. and L. B. Montefusco (2002). Radial basis functions for the multivariate interpolation of large scattered data sets. *Journal of Computational and Applied Mathematics 140*(1), 521–536.

Lee, J., P. K. Snyder, and P. F. Fisher (1992). Modeling the effect of data errors on feature extraction from digital elevation models. *Photogrammetric Engineering and Remote Sensing 58*(10), 1461–1467.

Leitner, M. and B. P. Buttenfield (2000). Guidelines for the display of attribute certainty. *Cartgraphy and Geographical Information Science 27*(1), 3–14.

Lemmens, M. J. P. M. (1997). Accurate height information from airborne laser-altimetry. In *Proceedings of the 1997 International Geoscience and Remote Sensing Symposium.*

Li, L. and P. Revesz (2004). Interpolation methods for spatio-temporal geographic data. *Computers, Environment and Urban Systems 28*, 201–227.

Li, Z. (1988). On the measure of digital terrain model accuracy. *Photogrammetric Record 12*(72), 873–877.

Li, Z. (1992). Variation of the accuracy of digital terrain models with sampling interval. *Photogrammetric Record 14*(79), 113–128.

Li, Z. (1993). Mathematical models of the accuracy of digital terrain model surfaces linearly constructed from square gridded data. *Photogrammetric Record 14*(82), 661–674.

Li, Z. (1994). A comparitive study of the accuracy of digital terrain models (DTMs) based on various data models. *ISPRS Journal of Photogrammetry and Remote Sensing 49*, 2–11.

Lichti, D. D. and S. J. Gordon (2004). Error propagation in directly georeferences terrestrial laser scanner point clouds for cultural heritage recording. In *Proceedings FIG Working Week 2004*, Number WSA2 Modelling and Visualization, Athens, Greece.

Ligda, M. G. H. (1963). In *Proceedings of Conference on Laser Technology*, Volume 1, San Diego, California, pp. 63–72.

Lillesand, T. M. and R. W. Kieffer (2000). *Remote Sensing and Image Interpretation.* John Wiley and Sons. 4th edition.

Lindenberger, J., P. Freiss, and S. Sizgoric (2003). Laser Ranging. Course Notes from International MSc Module.

Lloyd, C. D. and P. M. Atkinson (1998). Scale and the spatial structure of landform: optimising sampling strategies with geostatistics. In *Proceedings of the 3rd International Conference on GeoComputation*, University of Bristol. GeoComputation.

Lloyd, C. D. and P. M. Atkinson (2002). Deriving DSMs from LiDAR data with kriging. *International Journal of Remote Sensing 23*(12), 2519–2524.

Lohr, U. (1998). Laserscanning for DEM generation. In C. A. Brebbia and P. Pascalo (Eds.), *GIS technologies and their environmental applications*, pp. 243–249. Computational Mechanics Publications, Southampton, UK.

Longley, P. A., M. F. Goodchild, D. J. Maguire, and D. W. Rhind (2001). *Geographic Information Systems and Science.* John Wiley and Sons. Chichester.

Longley, P. A., M. F. Goodchild, D. J. Maguire, and D. W. Rhind (2005). *Geographic Information Systems and Science, 2nd Edition.* John Wiley and Sons. Chichester.

Longley, P. A. and V. Mesev (2003). Measuring urban morphology using remotely-sensed imagery. In V. Mesev (Ed.), *Remotely Sensed Cities*, pp. 163–184. Taylor and Francis, London.

Lopez, C. (1997). Locating some types of random errors in digital terrain models. *International Journal of Geographical Information Science 11*(7), 677–698.

Lyche, T., T. K. Nilssen, and R. Winther (2002). Preconditioned iterative methods for scattered data interpolation. *Advances in Computational Mathematics 17*(3), 237–256.

Maas, H.-G. (1999). The potential of height texture measures for the segmentation of airborne laser scanner data. In *Proceedings of the Fourth International Airborne Remote Sensing Conference and Exhibition*, 21st Canadian Symposium on Remote Sensing, Ottowa, Ontario, Canada.

Maas, H.-G. (2000). Least-squares matching with airborne laserscanning data in a TIN structure. *International Archives of Photogrammetry and Remote Sensing XXXIII*(Part B3), 548–555. Amsterdam.

Maas, H.-G. (2002). Methods for measuring height and planimetry discrepancies in airborne laserscanner data. *Photogrammetric Engineering and Remote Sensing 68*(9), 933–940.

Maas, H.-G. and G. Vosselman (1999). Two algorithms for extracting building models from raw laser altimetry data. *ISPRS Journal of Photogrammetry and Remote Sensing 54*, 153–63.

MacEachren, A. M. (1995). *How Maps Work.* The Guilford Press, London.

MacEachren, A. M. and J. V. Davidson (1987). Sampling and isometric mapping of continuous geographic surfaces. *The American Cartographer 14*(4), 299–320.

Mark, D. M. (1978). Concepts of data structure for digital terrain models. In *Proceedings Digital Terrain Model (DTM) Symposium*, St Louis, Missouri, pp. 24–31.

Martin, D. (1991). *Geographical information systems and their socioeconomic applications.* Routledge, London.

Martin, D. (1996). *Geographical information systems and their socioeconomic applications. 2nd Edition.* Routledge, London.

Martin, D. (2002). Census population surfaces. In P. Rees, D. Martin, and P. Williamson (Eds.), *The census data system*, pp. 139–148. Wiley, Chichester.

Martinez, W. L. and A. R. Martinez (2002). *Computational Statistics Handbook with MatLab.* Chapman and Hall.

Masaharu, H. and K. Ohtsubo (2002). A filtering method of airborne laser scanner data for complex terrain. In *International Archives of Photogrammetry and Remote Sensing*, Volume XXXIV 3B III, Graz, Austria, pp. 165–169.

Matikainen, L., J. Hyyppä, and H. Hyyppä (2003). Automatic detection of buildings from laser scanner data for map updating. In *International Archives of Photogrammetry and Remote Sensing*, Volume XXXIV 3/W13, pp. 218–224.

McClung, F. J. and R. W. Hellwarth (1962). Giant optical pulsations from ruby. *Journal of Applied Physics 33*, 828–829.

McCullagh, M. J. (1988). Terrain and surface modelling systems: theory and practice. *Photogrammetric Record 12*, 747–779.

McGranaghan, M. (1993). A cartographic view of spatial data quality. *Cartographic Perspectives 13*, 10–19.

Méaille, R. and L. Wald (1990). Using geographical information systems and satellite imagery within a numerical simulation of urban growth. *International Journal of Geographical Information Systems 4*, 445–456.

Measures, R. (1992). *Laser Remote Sensing: Fundamentals and Applications.* Krieger Publishing Company, Florida.

Mesev, V. (2003). *Remotely Sensed Cities.* Taylor and Francis, London.

Michalak, W. X. (1993). GIS in land use change analysis: Integration of remotely sensed data into GIS. *Applied Geography 13*, 28–44.

Monckton, C. (1994). An investigation into the spatial structure of error in digital elevation data. In M. Worboys (Ed.), *Innovations in GIS 1*, pp. 201–211. Taylor and Francis, London.

Moore, I. D., R. B. Grayson, and A. R. Ladson (1991). Digital terrain modelling: A review of hydrological, geomorphological, and biological applications. *Hydrological Processes 5*, 3–30.

Moore, P. (2004). Interpolation-based a posteriori error estimation for *hp*-refinement using first and 2nd derivative jumps. *Applied Numerical Mathematics*, 63–82.

Morgan, M. and A. Habib (2002). Interpolation of lidar data and automatic building extraction. In *ACSM-APRS 2002 Annual Conference Proceedings.*

Morgan, M. and K. Tempfli (2000). Automatic building extraction from airborne laser scanning data. *International Archives of Photogrammetry and Remote Sensing 33*(B3), 616–623. Amsterdam.

Neter, J., W. Wasserman, and G. A. Whitmore (1993). *Applied Statistics.* Allyn and Bacon, Boston, Massachusetts. Fourth Edition.

Nichol, J. E. (1988). Visualisation of urban surface temperatures derived from satellite images. *International Journal of Remote Sensing 19*, 1639–1650.

Nixon, M. and A. Aguado (2002). *Feature Extraction and Image Processing.* Newnes, Oxford.

Ogrosky, C. E. (1975). Population estimates from satellite imagery. *Photogrammetric Engineering and Remote Sensing 41*, 707–12.

Palm, M. and U. Lohr (2000). In *Surveying Coastal Erosion by Airborne Laserscanning*, pp. 238. ICG Publishing Ltd., London.

Palmer, T. C. and J. J. Shan (2001). Urban modelling from lidar data in an integrated GIS environment. In *Proceedings of ASPRS Conference*, Saint Louis, MO.

Pathan, S. K., S. V. C. Sastry, P. S. Dhinwa, M. Rao, M. K. Kumar, S. Patkar, V. N. Patkar, and V. N. Phatak (1993). Urban growth trend analysis using GIS techniques - a case study of the Bombay metropolitan region. *International Journal of Remote Sensing 19*, 823–854.

Patil, U., S. Ravan, and A. Kaushal (2003). GIS based air pollution surface modeling. *GIS Development August.*

Pfeifer, N. and D. R. R. D. Santos (2004). Least sqaures matching for comparison of digital terrain models and its application potential for the Brazilian models and the SRTM model. *Bol. Ciênc. Geod 10*(2), 177–191.

Pfeifer, N., P. Stadler, and C. Briese (2001a). Derivation of digital terrain models in the SCOP++ environment. In *OEEPE Workshop on Laser Scanning and Interferometric SAR for Digital Elevation Models*, Stockholm.

Pfeifer, N., P. Stadler, and C. Briese (2001b). Derivation of digital terrain models in the SCOP++ environment. In *OEEPE Workshop on Airborne Laserscanning and Interferometric SAR for Digital Elevation Models, Stockholm.*

Polgar, S. and S. A. Thomas (1991). *Introduction to Research in the Health Sciences.* Churchill Livingstone, London. 2nd edition.

Polidori, L., J. Chorowicz, and R. Guillande (1991). Description of terrain as a fractal surface. *Photogrammetric Engineering and Remote Sensing 57*(10), 1329–32.

Polit, D. F., C. T. Beck, and B. P. Hungler (2001). *Essentials of Nursing Research: Methods, Appraisal and Utilisation.* Williams and Wilkins, Philadelphia, Lippincolt. 5th Edition.

Priestnall, G. and R. Glover (1998). A control strategy for automated land use change detection: an integration of vector-based GIS, remote sensing and pattern recognition. In S. Carver (Ed.), *Innovations in GIS 5*, pp. 162–175. Taylor

and Francis, London.

Priestnall, G., J. Jaafar, and A. Duncan (2000). Extracting urban features from LiDAR surface models. *Computers, Environment and Urban Systems 24*, 65–78.

Quenouille, M. (1956). Notes on bias estimation. *Bimetrika 43*, 353–360.

Raber, G. (2003). The effect of lidar posting density on dem accuracy and flood extent delineation: A gis-simulation approach.

Ranzinger, M. and G. Gleixner (1995). Changing the city: data sets and applications for 3D urban planning. *GIS Europe 3*(95), 28–30.

Rees, D. G. (2001). *Essential Statistics*. Chapman and Hall.

Rees, W. G. (2000). The accuracy of digital elevation models interpolated to higher resolution. *International Journal of Remote Sensing 21*(1), 7–20.

Richman, A. (1999). Urban relief, CEO product development and marketing projects. URL. http://ewse.ceo.org/anonymous/vfs.pl?id=889117name=summary.pdf.

Roggero, M. (2001). Airborne laser scanning: clustering in raw data. In *International Archives of Photogrammetry and Remote Sensing*, Volume XXXIV, Annapolis, MD, pp. 227–232.

Roggero, M. (2002). Object segmentation with region growing and principal component analysis. In *ISPRS Commission III Symposium: Photogrammetric Computer Vision*, Graz, Austria, pp. 289–294.

Rottensteiner, F. and J. Jansa (2002). Automatic extraction of buildings from lidar data and aerial images. In *Proceedings of the Symposium on Geospatial Theory, Processing, and Applications*, Ottowa, Canada.

Sailor, D. J. and L. Lu (2004). A top-down methodology for developing diurnal and seasonal anthropogenic heating profiles for urban areas. *Atmpospheric Environment 38*(17), 2737–2748.

Sandwell, D. T. (1987). Biharmonic spline interpolation of GEOS-3 and SEASAT altimeter data. *Geophysical Research Letters 2*, 139–142.

Scaperdas, A. and R. Colvile (1999). Assessing the representativenessof monitoring data from an urban intersection site in central london. *Atmospheric Environment 33*, 661–674.

Schenk, T., S. Seo, and B. Csatho (2001). Accuracy study of airborne laser scanning data with photogrammetry. *International Archives of Photogrammetry and Remote Sensing XXXIV*(3/W4), 113–118.

Schinckler, W. and A. Thorpe (2001). Surface estimation based on lidar. In *Proceedings of the ASPRS Annual Conference*, St Louis, Missouri.

Schneider, M. (2004). Vague spatial data types. In *Proceedings of the International Symposium on Spatial Data Quality*, Volume 28a, Vienna.

Schofield, D., G. Hunter, and P. Kerr (2003). Virtually flooded: Visualising flood data in Morpeth. *Geomatics May/June*, 22–23.

Siegel, S. (1956). *Nonparametric Statistics for the Behavioral Sciences*. McGraw-Hill.

Sithole, G. (2003). Filtering strategy: Working towards reliability. In *International Archives of Photogrammetry and Remote Sensing*, Volume XXXIV 3A III, Graz, Austria, pp. 330–335.

Sithole, G. and G. Vosselman (2003). Report: ISPRS comparison of filters. In *International Archives of Photogrammetry and Remote Sensing*, Volume XXXIV 3/W13, Dresden, Germany, pp. 71–78.

Sithole, G. and G. Vosselman (2004). Experimental comparison of filter algorithms for bare-earth extraction from airborne laser scanning point clouds. *ISPRS Journal of Photogrammetry and Remote Sensing 59*, 85–101.

Skaloud, J. and K.-P. Schwarz (1998). Accurate orientation for airborne mapping systems. In *ISPRS Commission II WG1 Symposium on Data Integration: Systems and Techniques*, pp. 283–290.

Skidmore, A. K. (1990). Terrain position as mapped from a griddded digital elevation model. *International Journal of Geographic Information Systems 4*(1), 33–49.

Sohn, G. and I. Dowman (2002a). Building extraction using LiDAR DEMs and IKONOS images. *International Archives of Photogrammetry and Remote Sensing XXXIV*(Part 3/W13), 167–173. Dresden, Germany.

Sohn, G. and I. Dowman (2002b). Terrain surface reconstruction by the use of tetrahedral model with the MDL criterion. In *International Archives of Photogrammetry and Remote Sensing: ISPRS Commission III Symposium*, Volume XXXIV, Graz, Austria, pp. 336–344.

Song, J.-H., S.-H. Han, K. Yu, and Y.-I. Kim (2002). Assessing the possibility of land-cver classification using lidar intensity. In *Proceedings ISPRS Commission III Photogrammetric Computer Vision*, Volume 34, Graz, Austria.

Stehman, S. (1997). Selecting and interpreting measures of thematic classification accuracy. *Remote Sensing of the Environment 62*, 77–89.

Steinle, E. and T. Voegtle (2001). Effects of different laserscanning modes on the result of buildings recognition and reconstruction. In *International Archives of Photogrammetry and Remote Sensing*, Volume B3, Amsterdam, The Netherlands, pp. 858–865.

Steinle, E. and T. Vögtle (2000). Effects of different laser scanning modes on the results of building recognition and reconstruction. In *International Archives of Photogrammetry and Remote Sensing.*

Tao, C. V. and Y. Hu (2001). A review of post-processing algorithms for airborne LIDAR data. In *Proceedings of ASPRS conference*, St.Louis, Missouri.

Tatem, A. J., H. G. Lewis, P. M. Atkinson, and M. S. Nixon (2001). Land cover mapping from remotely sensed images at the sub-pixel scale using a hopfield neural network. *IEEE Transactions on Geoscience and Remote Sensing 39*, 781–796.

Tempfli, K. (1998). Urban 3D topologic data and texture by digital photogrammetry. In *Proceedings of ISPRS*, Tempa, Florida,USA. CD-ROM.

TerraSolid Oy (2004). *TerraScan User's Guide.* TerraSolid Oy.

Thiel, K.-H. and A. Wehr (1999). Operational data processing for imaging laser al-

timeter data. In *Proceedings of the Fouth International Airborne Remote Sensing Conference and Exhibition*, Volume 4, pp. 1320–1327.

Thiessen, A. H. (1911). Precipitation averages for large areas. *Monthly Weather Review 39*, 1082–1084.

Tomczak, M. (1998). Spatial interpolation and its uncertainty using automated anisotropic inverse distance weighting (IDW) - cross-validation/jackknife approach. *Journal of Geographic Information and Decision Analysis 2*(2), 18–30.

Toutin, T. and L. Gray (2000). State-of-the-art of elevation from satellite SAR data. *ISPRS Journal of Photogrammetry and Remote Sensing 55*, 13–33.

Turner, D. P., R. Dodson, and D. Marks (1996). Comparison of alternative spatial resolutions in the application of a spatially distributed biogeochemical model over complex terrain. *Ecological Modelling 90*, 53–67.

Vable, M. (2001). Minimizing the error near discontinuities in boundary element model. *Engineering Analysis with Boundary Elements 25*, 607–617.

vanEssen, R. F., G. W. A. Offermans, A. W. S. Helwig, and D. vanWilligen (1997). Reducing spatial decorrelation effects through multi-station dgnss. In *Proceedings of 26th Annual Technical Symposium of the International Loran Association*, Ottowa, Canada.

Voltz, M. and R. Webster (1990). A comparison of kriging, cubic splines and classification for predicting soil properties from sample information. *Journal of Soil Science 41*, 473–490.

Voogt, J. A. and T. R. Oke (2003). Thermal remote sensing of urban climates. *Remote Sensing of Environment 86*, 370–384.

Vosselman, G. (1999). Building reconstruction using planar faces in very high density height data. *International Archives of Photogrammetry and Remote Sensing 32*(Part 3/2W5), 87–92.

Vosselman, G. (2000). Slope based filtering of laser altimetry data. In *International Archives of Photogrammetry and Remote Sensing*, Volume XXXIII, Amsterdam, The Netherlands, pp. 935–942.

Vosselman, G. (2003). 3D reconstruction of roads and trees for city modelling. In *International Archives of Photogrammetry and Remote Sensing*, Volume XXXIV 3/W13, Dresden, Germany, pp. 231–236.

Vosselman, G. and S. Dijkman (2001). 3D building model reconstruction from point clouds and ground plans. *International Archives of Photogrammetry and Remote Sensing XXXIV*(3/W4), 37–43.

Vosselman, G., P. Kessels, and B. G. H. Gorte (2005). The utilisation for airborne laser scanning for three-dimensional mapping. *International Journal of Applied Earth Observation and Geoinformation 6*(3-4), 177–186.

Vosselman, G. and H.-G. Maas (2001). Adjustment and filtering of raw laser altimetry data. In *Proceedings of OEEPE workshop on airborne laser scanning and interferometric SAR for detailed digital elevation models.*

Wack, R. and A. Wimmer (2002). Digital terrain models from airborne laser scanner data - a grid based approach. In *Proceedings of IAPRS ISPRS Commission III Symposium*, Volume XXXIV, Graz, Austria, pp. 293–296.

Walsh, S. J., D. R. Lightfoot, and D. R. Butler (1987). Recognition and assessment of error in geographical information systems. *Photogrammetric Engineering and Remote Sensing 53*(10), 1423–1430.

Wang, Z. and T. Schenk (2000). Building extraction and reconstruction from lidar data. In *International Archives of Photogrammetry and Remote Sensing*, Volume XXXIII, Amsterdam.

Watson, D. F. (1992). *Contouring. A guide to the analysis and display of spatial data.* Pergamon.

Weber, C. and J. Hirsch (1991). Some urban measurements from spot data: urban life quality indices. *International Journal of Remote Sensing 13*, 3251–3261.

Wechsler, S. (2003). Perceptions of digital elevation model uncertainty by dem users. *URISA Journal 15*(2), 57–64.

Wehr, A. and U. Lohr (1999). Airborne laser scanning - an introduction and overview. *ISPRS Journal of Photogrammetry and Remote Sensing 54*, 68–82.

Weibel, R. and M. Heller (1990). A framework for digital terrain modelling. In *4th International Symposium on Spatial Data Handling*, Zurich, pp. 219–229.

Weibel, R. and M. Heller (1991). Digital terrain modelling. In D. J. Maguire, M. F. Goodchild, and D. W. Rhind (Eds.), *Geographical Information Systems: Principles and Applications*, pp. 269–297. Longman, London.

Weidner, U. (1995). Evaluation of building extraction from digital elevation models. Technical report, Institut für Photogrammetrie, Bonn.

Welch, R. (1982). Spatial resolution requirements for urban studies. *International Journal of Remote Sensing 3*, 139–146.

Wood, J. (1994). Visualising contour interpolation accuracy in digital elevation models. In H. Hearnshaw and D. J. Unwin (Eds.), *Visualization in Geographical Information Systems*, pp. 168–180. John Wiley and Sons, New York.

Wood, J. (1996). *The Geomorphological Characteristation of Digital Elevation Models*. Ph. D. thesis, University of Leicester.

Wood, J. and P. Fisher (1993). Assessing interpolation accuracy in elevation models. *IEEE Computer Graphics and Applications 13*(2), 48–56.

Xie, K., Y. Wu, X. Ma, Y. Liu, B. Liu, and R. Hessel (2003). Using contour lines to generate digital elevation models for steep slope areas: a case study of the loess plateau in north china. *Catena 54*, 161–171.

Yang, X. and T. Hodler (2000). Visual and statistical comparisons of surface modeling techniques for point-based environmental data. *Cartography and Geographical Information Science 27*(2), 165–175.

Yeh, A. G.-O. and X. Li (2001). Measurement and monitoring of urban sprawl. *Photogrammetric Engineering and Remote Sensing 67*, 83–90.

Zhang, J. and M. Goodchild (2002). *Uncertainty in Geographical Information*. Taylor and Francis, London.

Zimmerman, D., C. Pavlik, A. Ruggles, and M. P. Armstrong (1999). An experimental comparison of ordinary and universal krigin and inverse distance weighting. *Mathematical Geology 31*(4), 375–390.

Zinger, S., M. Nikolova, and H. Maitre (2002). 3d resampling for airborne laser data of urban areas. *International Archives of Photogrammetry and Remote Sensing XXXIV*(3A III), 418–423.

Zlatanova, S. (2000). *3D GIS for Urban Development.* ITC, Netherlands.